DOCUMENT RESUME

ED 065 325 SE 014 300

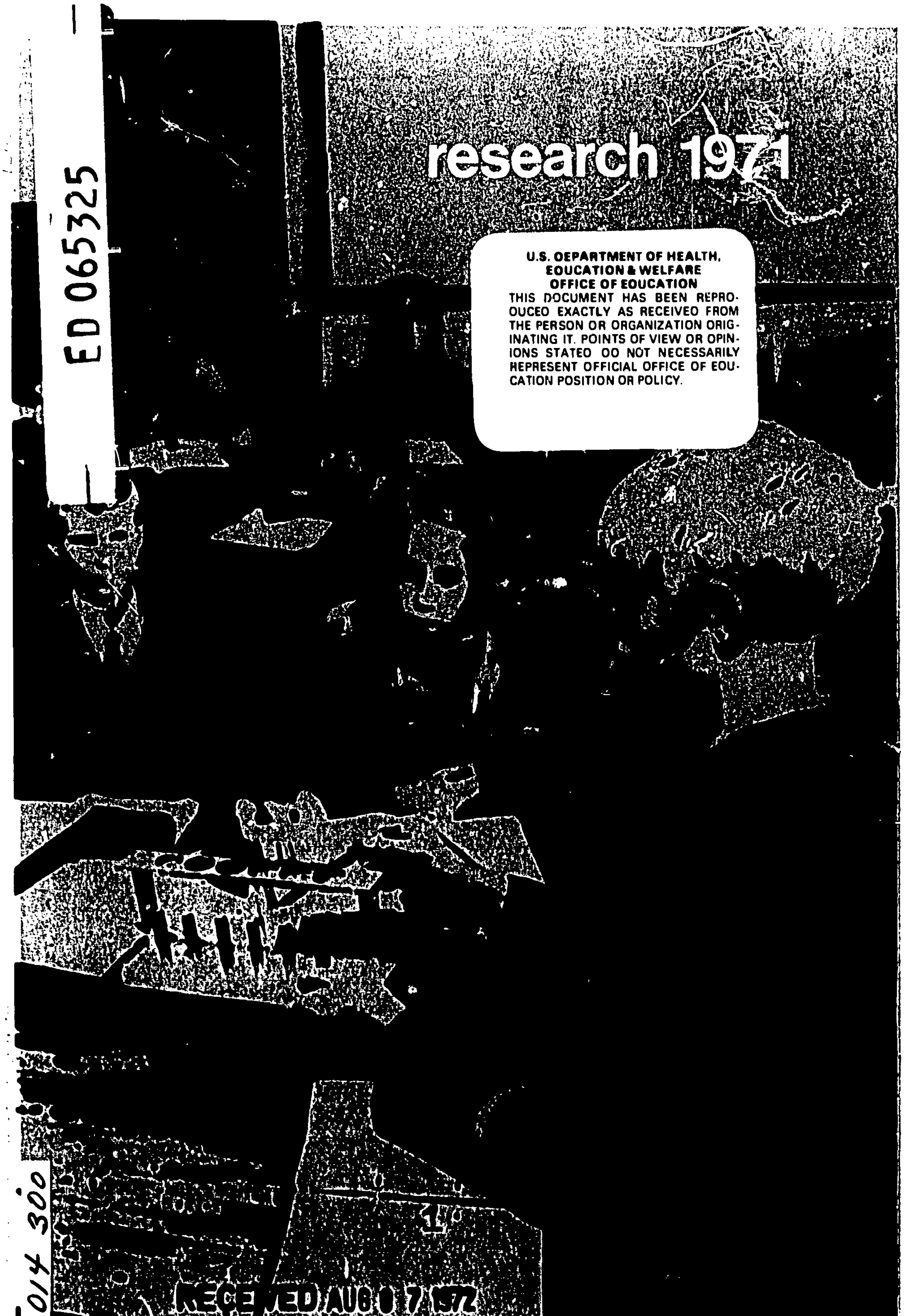
AUTHOR Tisher, R. P., Ed.
TITLE Research 1971.
INSTITUTION Australian Science Education Research Association, Clayton, Victoria.
PUB DATE 71
NOTE 146p.
AVAILABLE FROM University of Queensland, Department of Education, Brisbane, Q 4067, Australia ($3.50)

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS *Curriculum; Educational Research; *Evaluation; Instruction; Interaction Process Analysis; *Learning; *Science Education; Student Characteristics; *Teacher Education
IDENTIFIERS Australia

ABSTRACT

The eleven papers presented at the 1970 meeting of the Australian Science Education Research Association are arranged in five sections. The first two sections, "Countenance of Science Education Research" and "Cognitive Style," contain one paper each; the first, a review of research trends and the second, an experimental report. "Sequencing and Inquiry/Discovery Studies" contains three papers; one reporting the effects of varying amounts of guidance given students learning graphical skills, one comparing the effects of guided discovery and other methods of instruction in an elementary school science unit, and the third reporting correlations between components of inquiry methods and U.S. biology students' perceptions of instructional outcomes. Two studies using interaction analysis techniques, one in in-service teacher education and the other comparing the characteristics of students who exhibited different amounts of interaction with the teacher, are reported in "Micro-teaching and Interaction Studies." The final section, "Curriculum Evaluation," contains papers that report the plan for the formative evaluation of the Australian Science Education Project units, describe changes in content-free characteristics of students studying the different physics courses (e.g., cognitive preference, enjoyment of physics), outline problems of curriculum evaluation in developing countries, and assess a method of measuring scientific attitudes. (AL)

research 1971



THE AUSTRALIAN SCIENCE EDUCATION
RESEARCH ASSOCIATION

RESEARCH 1971

Edited by R.P. Tisher

A publication containing papers presented at the Associations' Annual Conference held in Sydney in May, 1971

Executive Secretary: Professor P. Fensham
Administrative Assistant: R.J. White
Faculty of Education,
Monash University,
Clayton. Victoria 3168

Publication printed by:
Australian Science Education Project,
Toorak. Victoria. 3142

Enquiries relating to "Research 1971" should be directed to:
The Editor, Faculty of Education, University of Queensland, Brisbane. 4067

CONTENTS

ERRATA : PREFACE

During the past two years informal meetings have been held by a group of persons engaged in research in science education. The "foundation" meeting occurred at Monash University, Melbourne, in 1970, and the second conference was organised at Macquarie University, Sydney, in May, 1971. At these last mentioned meetings, it was suggested that details of science education research in Australia should be disseminated more widely than had occurred in the past. Consequently, the idea of an Australian research report, containing copies of the papers read at the '71 conference was presented. This monograph is an indication that the idea was enthusiastically accepted! Certainly, this publication is a unique venture for Australian science education researchers. The document, too, is part of an embryonic project. A dream is that one day there will be an Australian journal reporting exclusively on research projects in science education. Hopefully, this present collection of papers is a precursor of a journal of research in science education. In the meantime, publications, such as the present one, appear to meet existing needs and to provide a means of disseminating information.

The first publication of the association is a modest one - certainly a group with limited funds and clientele must proceed cautiously - but the belief is that a fillip will be provided to science education research. The encouragement, co-operation and assistance of the Australian Science Education Project personnell is gratefully acknowledged. Without them this publication would have been impossible. However, the views, opinions, interpretations and implications expressed in the papers are not necessarily those of A.S.E.P. or of the Editor. They are those of the individual authors. Responsibility for research design, methodology and analysis, too, rests with each author. It is appropriate to note, however, that the reports, which evidence a wide range of interests, of necessity, vary in their levels of sophistication. This, however, is not a criticism but rather an indication and acknowledgement that a range and variety of studies are appropriate and essential in science education in the real world.

Ten of the twelve papers read at the Sydney conference appear in this publication. Although each report was prepared independently, as the manuscripts were collated, it seemed that some groupings could be made according to the common themes or emphases in some projects. Accordingly, the various reports have been placed into one of six categories designated, countenance of science education research, cognitive style, sequencing and inquiry, micro-teaching and interaction and curriculum evaluation. The responsibility for the grouping rests entirely with the editor, but hopefully, the categorization indicates the major areas being covered in current research projects.

This document is encouraging evidence of the burgeoning research in science education in Australia. It also presents a challenge to increase the volume of the research while maintaining its quality. Hopefully, the vigorous group of research workers will grow, as will the number of relevant, well-conceived projects. The future is indeed challenging and bright!

R.P. TISHER
University of Queensland

TRENDS IN SCIENCE EDUCATION: IMPLICATIONS FOR RESEARCH

R.P. Tisher

The objectives of this paper are to refer to some trends in science education and to suggest several implications for research. An attempt will be made to specify a number of research questions, and to indicate the types, or styles, of research which may be used to answer the questions. In addition, it is proposed to raise four important issues which are of concern to all science education researchers. It will not be possible, in the time available, to describe in detail some research designs. However, it may be possible, during the conference, for groups to meet to discuss the designs of future projects. This issue will be raised again later.

Research in science education is but one form of research which contributes to our understanding of the teaching-learning process. To emphasize the role this research plays and to provide a structure for the subsequent discussion, the trends in science education and implications for research will be discussed under four headings which are derived from the paradigm in figure 1. The paradigm shows some of the important elements in the teaching-learning process.

Elements of the Teaching-Learning Process

It is not appropriate here to describe the paradigm in great detail. However, there are some essential features which demand comment. First, an important element in the teaching-learning process is the interaction between the learner and other individuals or materials in the learning environment. These things or elements, designated S in the paradigm, may be other learners (or peers), teachers, and a variety of curriculum materials including books, apparatus, films and the like. Second, changes occur in the learner (from L to L') and the nature of the assessment of this change affects the nature of the of the interaction.

The nature of the interaction and of the assessment of growth will also affect and be affected by the nature of the classroom, school and community context. For example, physical facilities, administration patterns, and aims and values of the school administration and the community can affect the classroom interaction, and consequently, affect the nature of the learning. Trends in science education, too, have been brought about by changes in elements of the teaching-learning process and these trends will now be discussed under four headings derived from the paradigm. They are (a) Contexts, (b) Curricula, (c) The Learners, and (d) The Teacher. The categories are, not clearly, mutually exclusive ones, as will be evident in the subsequent discussion. Several trends, and the related research questions, can be placed under more than one heading.

Contexts

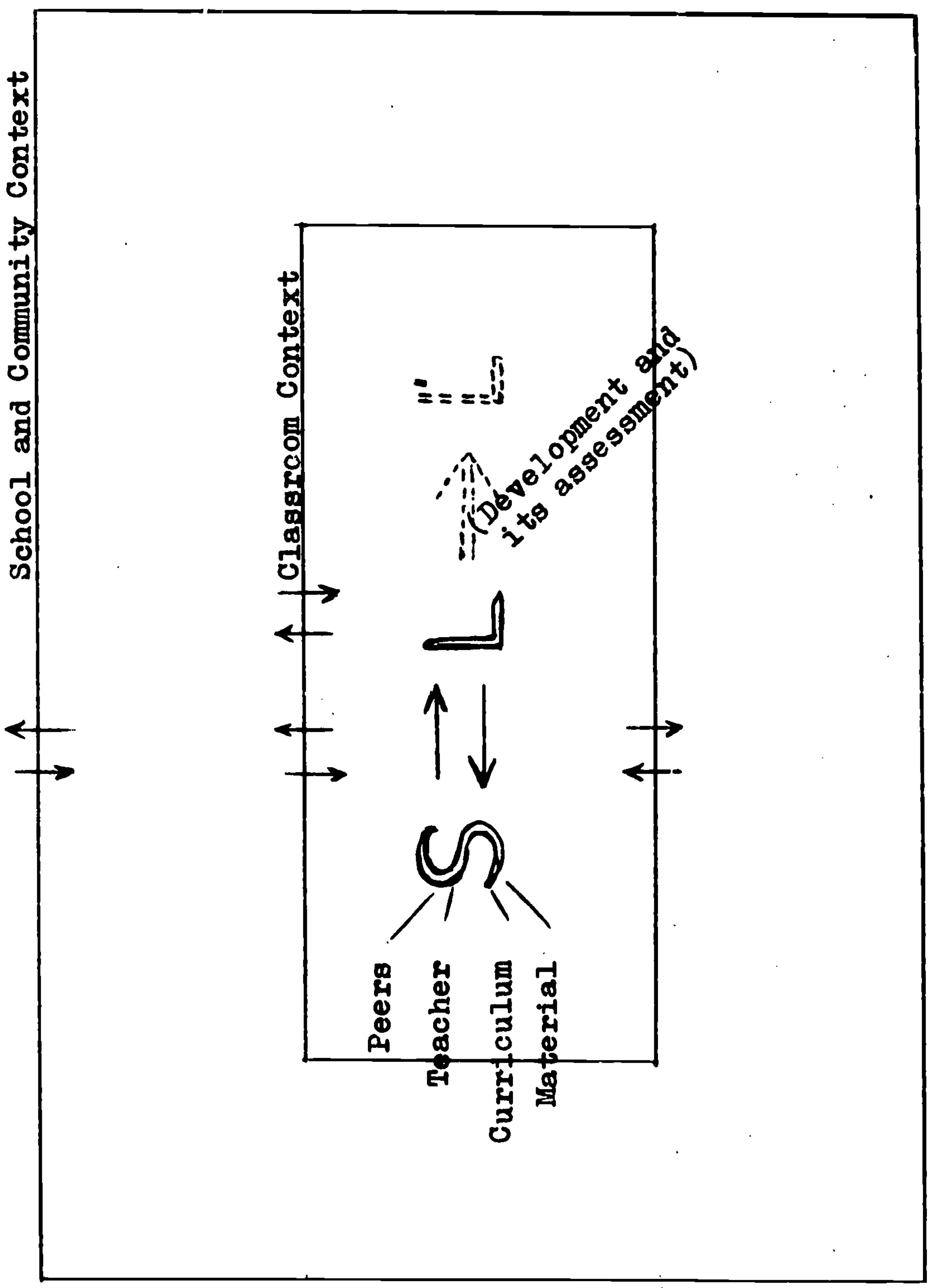
Increase in the size of the science class and participation in team-teaching are two trends that have been facilitated, in part, by changes in the educationists' and administrators' concept of the school and class context. The "open school", for example, is now an accepted feature in the U.S.A., and Australia, and consequences of the arrangement have been the creation of larger instructional groups for science and several science teachers co-operating in teaching a curriculum. Several traditionally designed schools, too, have attempted to emulate features of an open school. In Iowa, U.S.A., science teachers in one secondary school removed several walls from adjoining classrooms and co-operatively taught 90 children in the "new" science room. These teachers *believed* the arrangement fostered greater co-operation among themselves, was professionally and educationally stimulating, challenged the pupils and resulted in more efficient and effective learning. Certainly the children appear to be busy and hopefully, a great deal of learning takes place. Is this, in fact, the case? What are the effects of the specified changes? These questions appear to be legitimate ones for the research worker and the innovator. The answers must be obtained from carefully conducted studies in the *real world* (Atkin, 1968), not from tightly controlled and designed research projects. This is not meant to imply that the studies should be "sloppy". There is still the need to define problems clearly (Jacobsen, 1970), to give adequate attention to sampling, to provide a rationale for confidence limits (Skipper, et al, 1967) and to exercise care in interpretation.

Another trend, albeit a long-established, but not universally accepted one, is to provide science education in a room specially equipped for the purpose. All science lessons are held in the room and, hopefully, greater flexibility in teaching results. To what extent have science teachers capitalized on this newer arrangement? Has there been greater flexibility in teaching? What innovations have resulted? What factors inhibit the use of the room? These questions, too, are legitimate ones for the research worker.

It is generally accepted that the nature of administration in a school and the attitudes of principals inhibit many innovations in science education, e.g., individualized instruction, and open-ended inquiry and discovery. What factors and combinations of factors, on the other hand, enhance innovative practices, foster a zest for learning in science and produce a satisfied science staff? This is a very important question and one to which science education research workers have given little attention. There is a need for studies which examine the " climate ", tone, or health of a school, as it has been variously called, and the nature of science education in the school. There is, fortunately, a special climate scale (the Secondary School Climate Scale) designed for Australian conditions (Grassie, 1971) available for the purpose. The scale is used to provide data on hierarchy of authority in the school, professional leadership, cohesion and support, and organisational constraints. From the data judgments can be made on the professional nature of school and how this is affecting learning and innovation.

Figure I

ELEMENTS IN THE TEACHING-LEARNING PROCESS



Another issue, often ignored by science education research workers, is that of the affect of the "environmental press" on learning in science. Environmental press refers to the psychological impact of the environment; the environment being a complex stimulus. The press is reflected in the characteristic pressures, stresses, rewards and conformity-demanding influences of the environment. Generally, a measure of this press is obtained from pupil's perceptions. An Australian instrument (the Science Education Environment Index) (Genn, 1969) is especially suitable for a detailed study of the differences among various science education environments and will provide a measure of the impact of various innovations in science education. Hopefully, the introduction of new facilities and curricula have resulted in considerable changes in teaching strategies and in changes of both student and teacher press.

In passing, it is appropriate to refer to another development which, although not a trend in science education, is affecting the implementation of innovation in science teaching (or, so it is thought). The development is the in-service education of principals on the nature of contemporary science curricula. The in-service programmes include talks on the curricula, visits to schools using the new materials and lessons in which the principals work through sections of the course. The rationale for the in-service programmes is that the principals, as a result of their experiences, will become more sympathetic to the problems of the science teacher and enlightened in their understanding of the nature of the new courses. But, is this the case? Furthermore, are the effects of the in-service programmes short-term, or long-term ones? These are questions for some research workers but, at the same time, others could be developing and implementing a variety of in-service courses for principals, inspectors and education ministers!

Curricula

The well publicised changes in science education are those associated with science curricula. The various trends are documented in detail by Hurd (1969, 1970) and Hurd & Gallagher (1969), and Lockard (1968). Some curricula changes are associated with the substantive structure of the sciences, (Robinson, 1969; Smith, 1969; Fensham, 1969) key concepts which are essential for a scientifically literate person (Karplus, 1964), conceptual schemes (Shamos, 1968), the sequencing of learning tasks, self-pacing and individualized instruction (Hurd and Gallagher, 1969), the processes of science (Rutherford, 1964), the history and philosophy of science, the impact of science on society, and attempts to integrate the recognised fields of science (Klopfer, 1966). The changes are legion but from among the many it seems appropriate to highlight two. First, there is the increased provision for learning by the individual. The Elementary Science Study (Hurd and Gallagher, 1969) and the Intermediate Science Curriculum Study (Hurd, 1970) for example, are two contemporary programmes designed to foster individual learning by enquiry and discovery. In these curricula the teacher's role is to guide, and to clarify rather than to dominate and rigidly direct learning activities. The I.S.C.S. programme is interesting from another point of view too. This Junior Secondary course begins with experimental work in physics (electricity) followed by experimental work in chemistry. Biology and environmental studies do not appear until the later stages of the course. Surveys carried out by the curriculum designers reveal that pupils are quite excited and enthusiastic about the sequencing. In fact, they are less interested in some biology sections. The reasons for this lower interest are by no means clear but may be associated with the type of experimental work in physics, chemistry and biology. Be that as it may, there is also much to be discovered about the effects of different sequences in curricula.

Second, there are the attempts to produce curricula which are more relevant to pupils, and there are several answers to the question of "what is relevant"? Cossman (1969), for example, developed a science and culture course which he believed was relevant to pupils in a University laboratory school. Other workers have suggested science programmes dealing with themes associated with the problems facing mankind, e.g., increasing population, pollution, starvation and drug use and abuse. Some of these ideas are incorporated into sections of I.S.C.S. and the Environmental Studies Project (E.S.I., 1970). Certainly, the greater concern with the last mentioned themes reflects changes in the values and attitudes of society which is asking "what is science education's role in man's survival"? This question, too, is an important one for the curriculum developer and for the reseacher who is concerned with philosophical and historical issues. There is a great need for rigorous analyses of the assumptions underlying science curricula and for studies of the social and political factors which influence science curricula in different decades. A most potent criticism is that there have not been enough research projects in science education of the historical and philosophical types.

It is generally accepted that science curricula should be evaluated and, certainly, there is a need to examine in detail the consequences of curriculum implementation. This examination, or summative evaluation, has been undertaken, but generally few studies have gone beyond bivariate correlations, t-tests, or one-way analyses of variance (ANOVA) (Welch, 1969). Some investigators have used the semantic differential technique (Rothman, 1968; Walberg and Anderson 1968). Few studies have used multivariate procedures, e.g., multiple and canonical correlation, factor analysis, discriminant analysis and multivariate analyses of variance (MANOVA), yet our present computer facilities are quite capable of coping with multivariate programmes.

As well as summative evaluation there is the need for formative evaluation (Scriven, 1967; Grobman, 1968) in curriculum projects. This implies that research workers will continue to *develop* science programmes and to give greater attention to evaluation of the programmes at many stages in their development. Much formative and summative evaluation has, to date, been ill-conceived, inadequately designed and limited to studies of some changes in pupils and teachers. There is an urgent need to examine the congruence or dissonance between teachers' perception of curricula and those of the curriculum planners, to study classroom transactions (Grobman, 1968; Tisher, 1971a), to examine the effects on administrators, parents and other teachers (e.g., maths teachers), and to assess the affect of teacher preparation on curriculum implementation. Furthermore, research workers need to give greater attention to the nature of their sample and to the types of product measures used. For example, it may be more appropriate to use criterion-referenced measurement (Popham and Husek, 1969; Ward, 1970) rather than norm-referenced measurement. Many of the preceding ideas could well be included in the work of A.S.E.P. and in the examinations of curricula in N.S.W., as well as in projects in other States. The serious curriculum worker and researcher should, too, read the A.E.R.A. Monographs on Curriculum

Evaluation (Tyler et al, 1967) and Stake's views on " the countenance of educational evaluation " (Stake 1967).

The Learners

For decades educational psychologists have emphasized the nature of individual differences and the need to cater adequately for these in the teaching-learning process. Also, the benefits have been advanced of learning in small groups. It is no wonder, then, that recent developments in science education include a greater emphasis on learning in small groups and on learning tasks for the individual. The discipline-oriented courses such as P.S.S.C., CHEM study and B.S.C.S. while emphasizing the structure of physics, chemistry and biology respectfully, also provide ample opportunity for individuals to work on their own, or in groups. I.S.C.S., E.S.S., and the A.S.E.P. materials too, provide for individual and group work. However, the three last mentioned curricula are designed to cater for as great a clientele as possible - a clientele with varying abilities, prior experiences, educational sets and achievement orientations. How do individuals with differing abilities, sets and experiences, for example, cope with these courses? How effective are the group and individual learning situations for these pupils? Questions of these types are also important ones for the research worker. The answers can be obtained in well designed and controlled quasi-experimental studies (Campbell and Stanley, 1963).

The researcher who is concerned with the quality and quantity of learning cannot neglect to study the effectiveness of different strategies of teaching in interaction with pupils' abilities and prior experiences. In addition to the research on group and individual instruction there is the need to continue studies on the affects of inquiry on cognitive and affective behaviour of pupils (Raun and Butts, 1968), the use of visual organizers (Weisberg, 1970) and the effects of feedback (Zahorik, 1969; Power, 1969). The newer science curricula are constructed on assumptions that inquiry (Rowlands, 1969), organizers and feedback, for example, enhance learning for all pupils. But our knowledge of the role of each of these strategies is still meagre and more research on their effects is required. In fact, there is the need to study the effects of a variety of strategies of teaching. The burgeoning research on classroom interaction (Campbell, 1968; Nuthall, 1968b; Tisher, 1970a; Power, 1971) is a step towards the more effective studies of strategies but the task is especially challenging for the science education research worker who needs to study verbal and non-verbal behaviour in discovery, lecture, group and individual learning situations. Some classificatory schemes of classroom behaviour are partially suitable (Tisher, 1971b), e.g., S.C.A.S. (Science Curriculum Assessment System (Matthews and Phillips, 1968) and M.A.C.I. (Multi-dimensional Analysis of Classroom Interaction) (Honigman, 1970), and could well be modified and adapted for various projects. Be that as it may, the information obtained only describes the classroom *as it is*. There is a need, too, to study the effects of *new* teacher behaviours. One way of obtaining these is to use existing behaviours or strategies combined in ways not presently observed (Gage, 1966; Meux, 1967). It is believed (Gage, 1966; Meux, 1967) that these combinations may result in teaching strategies which are far more effective than those presently used by our best teachers. A study along the lines of this "combined strategies" proposal has been reported recently (Nuthall, 1968a), but in common with much other educational research it cannot be classified as an essentially multi-variate study. Certainly a failing in many of the learning studies in science education (Belanger, 1969) is that they are formulated in uni-variate rather than multi-variate terms. The literature contains many reports on the effects of teacher-personality and attitudes (Sadler, 1968; Mackay, 1969) cognitive style (Field and Cropley, 1969) open-mindedness and close-mindedness of pupils (Blankenship and Hay, 1968), differing structures in science content (Anderson, 1968), learner-supportiveness (Tisher, 1970b) and higher-cognitive questions (Tisher, 1970c) on pupil growth. Few studies consider all these factors together and examine their association with a *number* of pupil outcomes. It is imperative that many more research projects in science education be formulated in multi-variate terms and that where appropriate, researchers use techniques such as MANOVA, canonical correlation and discriminate analysis.

Another trend in science education, albeit a nascent one, is the attention being given to the disadvantaged learner and to the educationally uninvolved. That the concern with the uninvolved and disadvantaged is growing is evidenced by the presentation of several papers on the topic at the recent Nineteenth Annual Convention of the National Science Teachers Association, and by the appearance of several articles (Shoresman, 1964; Watson, 1967). The immediate research implications are that more workers should become involved in the *development* of science programmes which meet the needs of the disadvantaged and that, concurrently with the development, appropriate research studies should be initiated. In these studies attempts should be made to determine how to make science more relevent, how to increase pupils' motivation and their display of motivation in class, more coherent knowledge of readiness (Connell, et al, 1967), appropriate learning tasks and sequences of activities suitable for different kinds of learners.

It is appropriate here to refer again to the advent of self-paced programmes (e.g., I.S.C.S.) and to indicate that they present a challenge to contemporary techniques of examining in science. For decades there has been a stranglehold on educational practice as a consequence of of the ubiquitous use of tests and test items which *discriminate* among individuals. The techniques for constructing the discriminating items and tests are legitimate ones if an important purpose of the educational enterprise is to rank learners. However, this purpose is open to question. Moreover, some of the newer science programmes challenge it. In these programmes individuals proceed to a new learning task only when the previous one is completed ' satisfactorily" (according to certain specified criteria). Thus at each stage of entry to a new learning task pupils can be regarded as "successful". One difference among them is that some have completed more tasks than others. This does not necessarily imply that the former group is of A quality and the others, not. Furthermore, it is not appropriate to administer the same test to all students at a given stage of the year and to grade them on the basis of their scores. What, then, are appropriate procedures for reporting to pupils, parents and employers that relevant criteria have been met? What affect does this criterion-referenced testing (Ward, 1970) procedure have on the teaching-learning process? What are appropriate criterion-reference measures? These questions are important ones for the researcher who is interested in test-development, in evaluation of pupil outcomes and in enhancing community understanding of the teaching-learning process. Also, they are questions of great moment for science teachers who are adopting self-paced programmes.

The Teacher

It seems plausible to state that characteristics of the science teacher, e.g., intelligence, knowledge of subject matter, knowledge of education, learner-supportiveness, enthusiasm and personal commitment, will affect learning in pupils. That much research has yielded disappointedly low correlations between teacher characteristics and pupil learning is no surprise, especially as the majority of studies are of the univariate type. There is, then, a continuing need to identify those characteristics and competencies of the teacher, which in interaction with other factors, enhance learning in science. The need is particularly urgent at the moment, as there are many developments occurring in teacher-education. There are, for example, many burgeoning attempts to establish more realistic and effective pre-service and in-service courses (Popham, 1965), and to introduce new procedures into these programmes so that the eventual classroom performance of science teachers will change (Bruce, 1969). Micro-teaching is one of these newer procedures (Allen and Ryan, 1969) but its long-term effects have not been clearly established (Goldthwaite, 1968; Harris, et al, 1970). Science teachers and research workers could well co-operate in an extensive project to develop micro-teaching techniques and to study their long-term effectiveness.

At the same time there is much other developmental research to be conducted on the education of science teachers. There is the need to prepare self-paced and individualized pre-service programmes, to develop ancillary materials for the programmes and to engage in systematic try-outs in which feedback is used to guide and to improve the programme at each of its stages. There may be no .01 level findings from this developmental research, but there will be important changes in the nature of the science teacher's preparation - changes which can be judged in terms of criteria such as greater involvement of the teacher, zest for teaching and the like. It is *believed* the nature of the teacher's preparation affects the nature of science education in the schools, but more research is required to substantiate the belief and to indicate which factors are conducive to producing a satisfactory learning environment.

There is also a prevalent belief that in-service education affects the nature of science education. To what extent is this the case? What are the effects of programmes designed to upgrade the teacher's knowledge of science content, new strategies and innovative curricula? What types of in-service courses are the most effective in producing understanding of content? Are teachers' attitudes to science education changed as a result of the course? If so, how much? To what extent, if any, must a teacher's attitude change if he is to effectively teach a contemporary science curriculum? These questions are but a few of the many important ones to be answered in multi-variate studies. Furthermore, the questions must be related to a paradigm, or model, of teacher education, and there is, too, a need to formulate appropriate models. Some "model making" attempts have occurred in the U.S.A. (Karplus, 1964; Smith, 1969) and they may provide guidelines for the Australian science research worker.

Research in Science Education: Critical Issues

The preceding discussion contains an outline of some of the trends in science education and a number of questions for research. There was, also, an underlying assumption (which influenced several explicit suggestions) that there are four paramount types, or styles (Atkin, 1968) of research in science education. These types are (Joyce, 1968):

(a) philosophical research; that is, research which involves the analysis of assumptions underlying actions, the delineation of problems and an analysis of possible consequences of proposed decisions and actions.

(b) historical and comparative research; that is, research which involves studies of policy decisions, their implementation and consequence, analytical and critical studies of factors which have shaped educational developments in the past, and comparative - analytic studies of science education in other countries.

(c) empirical research; that is, research which involves the collection of data in specifically "designed" situations. The data may concern the behaviour of teachers or pupils, characteristics of teachers or pupils, and learning outcomes. Often, the data are used to accept or reject hypotheses.

(d) developmental research; that is, research which involves the preparation of educational programmes and materials, and systematic try-outs in which feedback is obtained to be used for improvements.

It is not appropriate here to elaborate on each of these types; detailed discussions appear elsewhere (Atkin, 1968; Joyce, 1968). It is appropriate however, to present four critical issues of concern to all researchers whether they are engaged in one, or several, types of science education research. The issues are (a) the definition of problems, (b) the use of null hypotheses, (c) initiation of co-operative research projects and (d) the training of research workers.

(a) The definition of problems

Defining a problem is probably the most critical, difficult and frustrating aspect of research. Yet, it is most essential. To conduct effective and efficient research, the problem must be seen clearly and defined precisely. A potent criticism of some research in science education is that the research problems have not been sufficietly clear and precise. Consequently, on occasions illegitimate procedures and techniques for analysis have been used. It behoves the researcher to give much attention and care to the definition of problems.

Similarly, the research worker must pay attention to his research design and the theoretical framework of his study. Certainly there is a need in science education research for more rigorous attention to the theoretical underpinnings (Tyler, 1968a; Tyler, 1968b) and consequences of the research.

(b) The use of null hypotheses

The ubiquitous use of the null hypothesis cannot remain unchallenged. The procedure is actually a disguised form of confirmation (Eastwood, 1967) and is open to many forms of abuse. The serious researcher should read widely on the topic (Eastwood, 1967). Furthermore, more hypotheses could be phrased as directional ones and one-tailed tests used to determine statistical significance. In passing, it is also appropriate to note there is nothing sacred about .01 or .05 levels of significance for hypothesis testing (Labovitz, 1968; Skipper, et al, 1967). At times, too, it may be appropriate to adopt several different levels providing, of course, a rationale is clearly presented.

(c) Initiation of co-operative research projects

There is a regrettable tendency in research in science education to develop and execute projects in isolation from one another. It is naive to expect that great quantities of these studies will significantly affect educational theory and practice. Education, including science education, is complex and subtle, and needs to be studied from many perspectives which are linked to each other. There is, I aver, a need for large-scale, co-operative, "umbrella-type" research projects in which many facets are studied at the one time but all projects are guided by, and contribute understanding to, the same model of science education.

(d) The training of research workers

The majority of research workers in science education slowly develop an expertise in educational research, which differs in many respects to research in science. There is a need to accelerate the development of the expertise and to acquaint researchers, at an early stage, with the issues associated with definition of problems, developing a theoretical framework, formulating hypotheses, sampling, data gathering and analysis, research designs, nature of historiography and comparative research, philosophy of pragmatism, non-parametric statistics and multi-variate techniques, and the like. Perhaps several future conferences could be arranged with the express purpose to develop expertise in science education research.

Research in science education places an equal, if not greater, demand on its researchers as does biology, physics, chemistry, animal husbandry, geology, geophysics, and agriculture. Certainly, systematic, well-conceived, well-executed, and courageous science education research is required. This implies that there will be courageous, knowledgeable, well-trained experts to conduct the research.

Concluding Comment

An attempt has been made, in this paper, to relate some trends in science education and some research questions to a paradigm of the teaching-learning process. Also, there were some provocative comments on the nature of research in science education and on four critical issues of concern to researchers. The discussion was intended to be a stimulus to greater research activity. That there is already burgeoning research in science education in Australia is encouraging. Certainly, there is a challenge to increase the volume of this research, while maintaining its quality. Hopefully, the vigorous group of research workers will grow, and it will continue to generate relevant, well-conceived projects. Hopefully, too, much valid research evidence will accrue to influence educational theory and practice.

REFERENCES

Allen, D. & Ryan, K. (1969). *Microteaching*, Reading, Addison-Wesley.

Anderson, O.R. (1968) "The Effects of Varying Structure in Science Content on the Acquisition of Science Knowledge", *Journal of Research in Science Teaching, 5*, 361-364.

Atkin, J.M. (1968) "Research Styles in Science Education", *Journal of Research in Science Teaching, 5*, 338-345.

Belanger, M. (1969) "Learning Studies in Science Education", *Review of Educational Research, 39*, 377-395.

Blankenship, J.W. & Hay, W.K. (1968) "An Analysis of the Relationship Between Open- and Closed-Mindedness and Capacity for Independent Thought and Action", *Journal of Research in Science Teaching, 5*, 69-72.

Bruce, M.H. (1969) "Teacher Education in Science", *Review of Educational Research, 39*, 415-427.

Campbell, W.J. (1968) "Studies of Teaching. I. Classroom Practices", *New Zealand Journal of Educational Studies, 3*, 97-124.

Campbell, D.J. & Stanley, J.C. (1963) "Experimental and Quasi-Experimental Designs for Research on Teaching" - Gage, N.L. (Editor), *Handbook of Research on Teaching*, Chicago, Rand McNally.

Connell, W. et al (1967) "Readiness for Learning" in *The Foundations of Education*, Sydney, Ian Novak.

Cossman, G.W. (1969) "The Effects of a course in Science and Culture for Secondary School Students", *Journal of Research in Science Teaching*, 6, 274-283.

Eastwood, G.R. (1967) "Confirmation and Null Hypotheses", *Educational Theory, 17*, 120-126.

E.S.I. (1970) *Environmental Studies Newsletter*, E.S.I., Environmental Studies Project, Boulder, Colorado.

Field, T.W. & Cropley, A.J. (1968) "Structures of thought among Senior Students: Science and Non-Science", *Australian Science Teachers' Journal, 14*, No. 2, 27-32.

Field, T.W. & Cropley, A.J. (1969) "Cognitive Style and Science Achievement", *Journal of Research in Science Teaching*, 6, 2-10.

Fensham, P.J. (1969) "Curricula and Materials in the Learning of Science", *The South Australian Science Teachers' Journal*, December, 34-40.

Gage, N.L. (1966) "Research on Cognitive Aspects of Teaching" in *The Way Teaching Is*, Report of the Seminar on Teaching. Association for Supervision and Curriculum Development, Washington.

Genn, J. (1969) *Chemistry scholars in school and University contexts - A study of transition from school to university*, Unpublished Ph.D. Thesis, University of Queensland.

Goldthwaite, D.T. (1968) *A study of Micro-Teaching in Pre-service Education of Science Teachers*, Unpublished Ph.D. Thesis, Ohio State University, Columbus.

Grassie, McC. (1971) *Relationships between Organizational Climate and Structure and the Professional Role Orientation, Position Held and Satisfaction of Teachers*, Unpublished Ph.D. Thesis, University of Queensland.

Grobman, H. (1968) *Evaluation Activities of Curriculum Projects*, AERA Monograph Series on Curriculum Evaluation, Chicago, Rand McNally.

Harris, W., Lee, V., & Pigge, F. (1970) "Effectiveness of Micro-Teaching Experiences in Elementary Science Methods Classes", *Journal of Research in Science Teaching, 7,* 31-33.

Honigman, F. (1970) "Multi-Dimensional Analysis of Classroom Interaction" in Simon A. & Boyer, E. (Eds.) *Mirrors for Behavior II*, Special Issue of Classroom Interaction Newsletter, Philadelphia, Research for Better School, Inc.

Hurd, P.D. (1969) *New Directions in Teaching Secondary School Science*, Chicago, Rand McNally.

Hurd, P.D. (1970) *New Curriculum Perspectives for Junior High School Science*, Belmont, Wadsworth.

Hurd, P.D. & Gallagher, J.C. (1969) *New Directions in Elementary Science Teaching*, Belmont, Wadsworth.

Jacobsen, W.J. (1970) "Approaches to Science Education Research: Analysis and Criticism", *Journal of Research in Science Teaching, 7,* 217-225.

Joyce, B.R. (1968) *The Teacher Innovator: A Program to Prepare Teachers*, U.S. Department of Health, Education and Welfare, Office of Education, Project No. 8-9019, Teachers' College.

Karplus, R. (1964) "The Science Curriculum Improvement Study - Report to the Piaget Conference" *Journal of Research in Science Teaching, 2,* 236-240.

Klopfer, L. (1966) "Integrated Science for the Secondary School: Process, Progress and Prospects", *Science Teacher, 33,* 27-31.

Labovitz, S. (1968) "Criteria for Selecting a Significance Level: A note on the Sacredness of .05", *The American Sociologist, 3,* 220-222.

Lockard, J.D. (Ed.) (1968) *Sixth Report of the International Clearinghouse on Science and Mathematics Curricular Developments*, College Park, University of Madison.

Mackay, L.D. (1969) "Teacher Characteristics Affecting Student Changes in P.S.S.C.", *Australian Science Teachers' Journal, 16,* No. 3, 74-77.

Matthews, C. & Phillips, D. (1968) *Handbook for the Application of the Science Curriculum Assessment System*, Department of Science Education, Florida State University.

Meux, M. (1967) "Studies of Learning in the School Setting", *Review of Educational Research, 37,* 539-562.

Nuthall, G.A. (1968a) "An Experimental Comparison of Alternative Strategies for Teaching Concepts", *American Educational Research Journal, 5,* 561-584.

Nuthall, G.A. (1968b) "Studies of Teaching, II. Types of Research on Teaching", *New Zealand Journal of Educational Studies, 3,* 125-147.

Popham, J. (1965) "An Experimental Attempt to Modify the Instructional Behavior of Student Teachers", *Journal of Teacher Education, 16,* 461-465.

Popham, W.J. & Husek, T.R. (1969) "Implications of Criterion-Reference Measurement", *Journal of Educational Measurement, 6,* 1-9.

Power, C.N. (1969) "The Question and Answer Game". *The South Australian Science Teachers' Journal.* December, 30-33.

Power, C.N. (1971) *The Effects of Feedback and Communication Patterns on Student Cognitive Growth in Science.* unpublished Ph.D. Thesis, in preparation, University of Queensland.

Raun, C. & Butts, D. (1968) "The Relationship between the Strategies of Inquiry in Science and Student Cognitive and Affective Behavioral Change", *Journal of Research in Science Teaching, 5,* 261-268.

Robinson, J.T. (1969) "Philosophical and Historical Bases of Science Teaching", *Review of Educational Research, 39,* 459-471.

Rothman, A. (1968) "The Factor Analysis of a Science-Related Semantic Differential Instrument", *Journal of Educational Measurement, 5,* 145-149.

Rowlands, E.C. (1969) "Fact and Fiction in Learning by Discovery", *Australian Science Teachers' Journal, 15,* No. 1, 5-10.

Rutherford, F.J. (1964) "The Role of Inquiry in Science Teaching", *Journal of Research in Science Teaching, 2,* 80-84.

Sadler, P. (1968) "Teacher Personality Characteristics and Attitudes Concerning P.S.S.C. Physics", *Journal of Research in Science Teaching, 5, 28-29.*

Scriven, M. (1967) "The Methodology of Evaluation", in Tyler, R., Gagne, R. and Scriven, M. *Perspectives of Curriculum Evaluation,* AERA Monograph Series on Curriculum Evaluation, No. 1, Chicago, Rand McNally.

Shamos, M.H. (1968) "The Role of Major Conceptual Schemes in Science Education", *Science Teacher, 33,* 27-30.

Shoresman, P.B. (1964) "Comment and Criticism; Grade Placement - Need for a Shift in Emphasis?", *Journal of Research in Science Teaching, 2,* 77-79.

Skipper, J.K., Guenther, A.L. & Nass, G. (1967) "The Sacredness of .05: A Note Concerning the Uses of Statistical Levels of Significance in Social Science", *The American Sociologist, 2,* 16-18.

Smith, B.O. (Ed.) (1969) *Teachers for the Real World,* American Association of Colleges for Teacher Education, Washington.

Smith, H.A. (1969) "Curriculum Development and Instructional Materials", *Review of Educational Research, 39,* 397-413.

Stake, R.E. (1967) "The Countenance of Educational Evaluation", *Teachers College Record, 68,* 523-540.

Tisher, R.P. (1970a) *A Study of Verbal Interaction in Science Classes and its Association with Pupils' Understanding in Science,* University of Queensland Papers, Vol. 1, No. 9, University of Queensland Press.

Tisher, R.P. (1970b) "The Nature of Verbal Discourse in Classrooms" in Campbell, W.J. (Ed.), *Scholars in Context,* Sydney, John Wiley.

Tisher, R.P. (1970c) "Association between Verbal Discourse and Pupils' Understanding in Science", in Campbell, W.J. (Ed.), *Scholars in Context,* Sydney, John Wiley.

Tisher, R.P. (1971a) "A Study of Teaching: An aspect of Curriculum Evaluation", *Curriculum Theory Network,* Ontario Institute for Studies in Education, (in press).

Tisher, R.P. (1971b) "Analyses of Classroom Behaviour : Their Relevance for Teachers", *Australian Science Teacher Journal, 17,* No. 2, (in press).

Tyler, R.W. (1968a) "Resources, Models, and Theory in the Improvement of Research in Science Education", *Journal of Research in Science Teaching, 5,* 43-51.

Tyler, R.W. (1968b) "Analysis of Strengths and Weaknesses in Current Research in Science Education", *Journal of Research in Science Teaching, 5,* 52-63.

Tyler, R., Gagne, R. & Scriven, M. (1967) *Perspectives of Curriculum Evaluation,* AERA Monograph Series on Curriculum Evaluation, No. 1, Chicago, Rand McNally.

Walberg, H. & Anderson, G. (1968) "Classroom Climate and Individual Learning", *Journal of Educational Psychology, 59,* 414-419.

Ward, J. (1970) "On the concept of Criterion-Reference Measurement", *British Journal of Educational Psychology, 40,* 314-323.

Watson, F.G. (1967) "Teaching Science to the Average Pupil", *The Science Teacher, 34,* 24-26.

Weisberg, J.S. (1970) "The use of Visual Advance Organizers for Learning Earth Science Concepts", *Journal of Research in Science Teaching, 7,* 161-165.

Welch, W.W. (1969) "Curriculum Evaluation", *Review of Educational Research, 39,* 429-443.

Zahorik, J. (1969) "Teacher Verbal Feedback", *Classroom Interaction Newsletter, 4,* No. 2, 34-41.

SOME DIMENSIONS OF A SCIENCE COGNITIVE STYLE AND THEIR RELEVANCE FOR THE CLASSROOM

T.W. Field

An important trend in recent curriculum developments in school science has been the increasing emphasis placed on such matters as concept formation, systematic thinking, inquiry processes, independent student investigations and the like. The teacher's role, of necessity, has had to become adapted to these new approaches. In the science class of today, or perhaps tomorrow, the teacher is no longer mainly an information giver, but rather his function is to guide and stimulate information processing by his pupils. Programmes such as Harvard Project Physics, Intermediate Science Curriculum Study, Web of Life, Chem Study, and the developing Australian Science Education Project materials, all involve this change of teaching strategy to varying extents. A common feature of these programmes is increased interaction between the teacher and individual pupils, and the effectiveness of this interaction must depend to a large extent on the teacher's perception of how pupil understandings develop. But how do children go about learning science? What, if any, types of thinking characterise the successful science student - the student who can process science information efficiently? Are there similarities in the thinking abilities of competent science students of different ages?

This paper reports some findings from an investigation of a number of thinking abilities among male science students at three different age levels: Form VI, and second-year university. At each age level, the pattern of thinking, or *cognitive style*, of a group judged to be highly orientated towards science was compared with that of a group of low science orientation. In addition, the comparison at the university level included a third group - engineering students - considered to be of middle-level science orientation. It was then possible to examine the cognitive styles of the science specialists (high science orientation groups) at the three age levels, and to attempt to define some dimensions of a 'science cognitive style'. In fact, there were quite marked similarities in the patterns of thinking abilities of the three science specialist groups, and these patterns were distinctly different from those of the other groups involved in the study. The specific details of this 'science cognitive style' are discussed below together with some possible implications of the findings for the classroom. But first, the concept of cognitive style, especially as applied here, will be amplified and the precise nature of the thinking abilities examined will be defined.

The Concept of Cognitive Style

What is a cognitive style? As mentioned above it can be thought of as a pattern of thinking, but it is more than this - it has a rather special personal quality about it. The term is mainly used in cognitive psychology to describe some of the characteristic ways in which a person thinks. Research studies have shown that an individual's thinking is marked by relatively consistent preferences, by favoured ways of approaching problems, or of dealing with information, and the concept of cognitive style embodies all these things. As Witkin, a notable worker in this area, defines it, cognitive style is " . . . the characteristic, self-consistent ways of functioning shown by a person in the cognitive (i.e. intellectual) sphere . . . " (1964, p.172). From this point of view, then, individual differences in ways of thinking can be understood in terms of different cognitive styles that take account of each person's pattern of relative strengths and weaknesses over a range of particular thinking abilities. If this is so, then it is also possible that one type of cognitive style may be more appropriate than others for study and learning in a particular subject area. That is, study in science may call for a cognitive style that is not really suited to the study of languages or history. Even within the sciences it is possible that specialisation in a particular area, such as chemistry, or physics, or biology, especially at an advanced level of study, may be associated with a cognitive style that is specific to that area. Though an interesting speculation this suggestion was beyond the scope of the present investigation which was concerned with identifying cognitive styles among senior high school and junior undergraduate students only.

Cognitive Style Measures

What types of basic abilities might be important aspects of the thinking of students who are highly orientated towards the study of science rather than other academic disciplines? Published research findings and the results of some preliminary studies suggested six major thinking variables, or dimensions of cognitive style; each of these is described briefly below together with some details of how it was measured.

Conceptual Differentiation This ability is concerned with one of the basic strategies of concept formation which can be thought of as a searching process. Concept formation involves looking for similarities and differences in a collection of information so that the information can be differentiated into groups (concepts) each of which contains information that shares some common property. Conceptual differentiation, then, is the ability to identify these common properties and was measured in this study as the number of groups formed by subjects on a sorting test adapted from the work of Gardner *et al* (1960). Actually two tests were used - one involving objects and the other photographs of people's faces - to follow up suggestions that science orientation is associated with a preference for dealing with things and abstractions rather than with people (Bush, 1969).

Conceptual Preference The formation of concepts is also likely to be influenced by individual preferences. One person may tend to group information in terms of one type of attribute while another person may favour a quite different type of attribute as the basis of his groupings. The reasons given by subjects for groupings on the two sorting tests were analysed to assess different types of preferences. The analysis was based on a classification system developed by Kagan *et al* (1963) and adapted by Wallach and Kogan (1965). Reasons for grouping objects were classed as analytic-descriptive, categorical-inferential, or relational, and those for grouping faces as physical, role, or pyschological.

21

Originality As measured here originality is concerned with the ability to think of remote consequences, or outcomes of an event. In a sense it involves an awareness of unusual possibilities or results of a given situation, or even an awareness of connections between apparently unrelated information. It was assessed by means of Guilford's *Consequences* test (Christenson *et al*, 1962).

Ideational Fluency Whereas originality is a measure of the quality of responses, ideational fluency simply takes account of their number or quantity. It is the ability to think of a large number of possible consequences of an event. *Consequences* also provided this measure.

Flexibility As the name implies, this type of thinking involves the ability to change frames of reference, and to see alternate ways of relating data, or of applying information. Flexibility is another important quality underlying concept formation and Guilford's *Alternate Uses* test (Wilson *et al*, 1960) was used to measure it.

Category Width When it comes to ordering or arranging information some people appear to be able to think in terms of broad, inclusive categories whereas others appear to use narrow exclusive ones. It is possible that this quality of thinking is also related to concept formation which can be thought of as involving the integration and organisation of information into broad inclusive categories. Pettigrew's (1958) *Category Width Scale* was used to obtain measures on this variable.

These, then, are the variables, the types of thinking abilities or characteristics, that were selected for study. No attempt can be made here to present the detailed theoretical justification for their selection, but the thumb-nail sketch given for each one should provide some appreciation of their relevance for successful thinking in the field of science. In fact, all the variables have parallels in the general psychological picture of mature scientists that has developed from a large number of research studies (Barron, 1965, provides an interesting summary). Some of the variables have been considered separately with various school age groups but, to date, there has been no consistent research effort to determine how the overall system of variables is related to science orientation at different age levels and degrees of specialisation.

The Sample

The sample was drawn from three metropolitan boys' high schools and from three universities, two metropolitan and one rural. The three hundred and fifty seven subjects were distributed across the three age levels and science orientation groups as shown in Table 1.

The tests were administered to form groups at each of the high schools but subjects were only retained in the sample if they attained a pass at the level for which they entered at the School Certificate or Higher School Certificate examinations. At the university level it was necessary to conduct the testing within the separate subject areas, and to accept course membership as the criterion of science orientation.

TABLE 1

Composition of Sample by Science Orientation
Grouping and Age Level

Level	Science Orientation Group (Course)	N	Total
Form IV	High (Advanced)	91	
	Low (Ordinary)	57	148
Form VI	High (1st Level/2F)	64	
	Low (3rd Level)	55	119
University II	High (Physics + Chem.)	35	
	Middle (Elect. Eng.)	28	
	Low (History)	27	90
	Total	(357)	357

Cognitive Style Findings

As mentioned above, the cognitive styles of the groups were compared at each age level separately. The five tests provided fifteen experimental variables and it was the pattern of scores on these variables that was defined as a cognitive style. Differences in the score patterns of the groups were investigated by means of multiple discriminant analysis (Veldman, 1967). In effect, this type of analysis allowed a comparison of the groups on all fifteen variables at the same time, and provided a description of differences in cognitive style in terms of sets of variables operating together. These sets of variables are defined by discriminant functions. The discriminant analysis in this study gave rise to one discriminant function at each of the school levels, and two at the university level, since the maximum number of functions that can be obtained is equal to either one less than the number of groups, or the number of variables, whichever is the smaller.

The first, the most important, finding was that within each age level the groups had distinctly different cognitive styles as can be seen in the three parts of Figure 1. The plots of the group centroids in this figure show the overall differences among the groups on the fifteen variables combined. In each case these differences were highly significant.

To determine the nature of the differences in cognitive styles it is necessary to examine the contributions that each variable makes to the various discriminant functions. There is not the space in this paper to present the detailed description of each of the three analyses. Instead, the findings for the university groups will be discussed and similarities and differences in the school level findings mentioned only briefly.

Figure 1

Group Discriminations

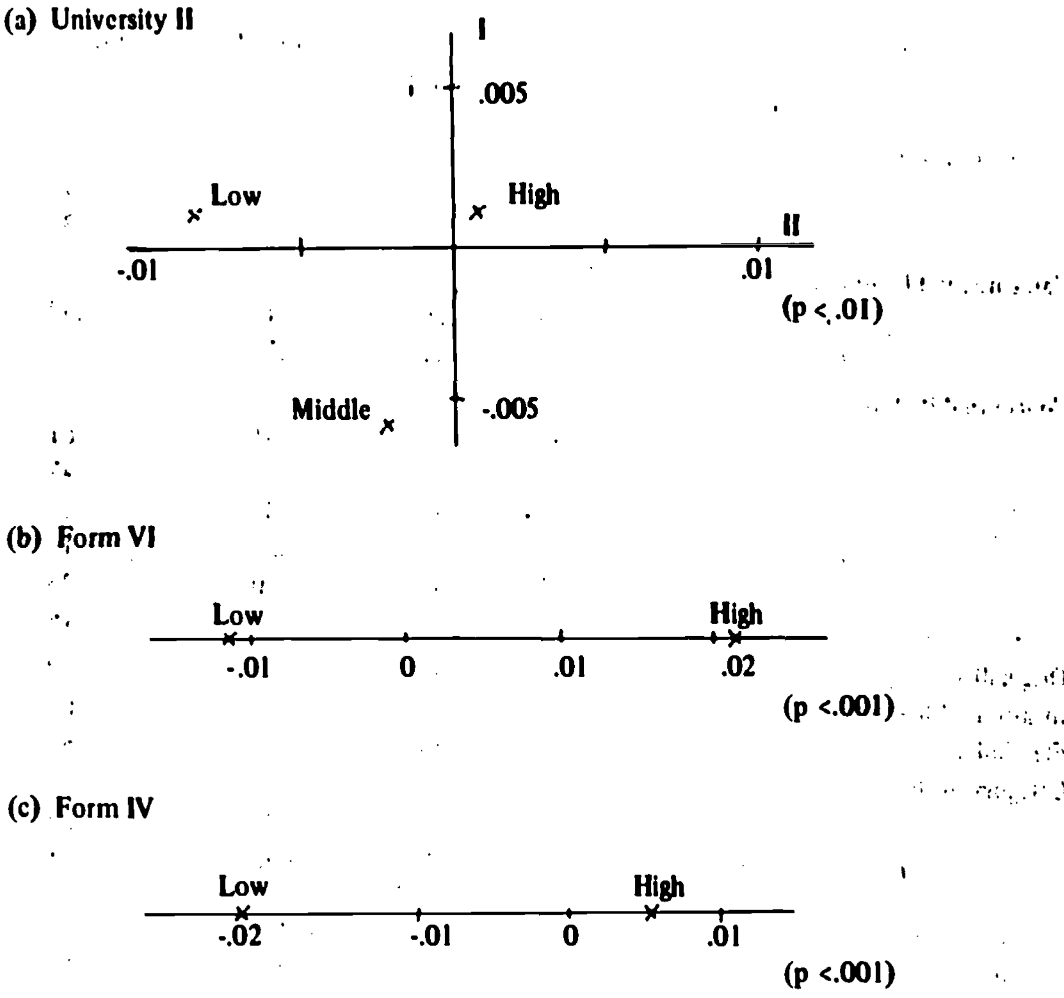


Table 2 shows the relative contributions made by the variables to the discriminant functions in the analyses at each of the three age levels. The second function obtained in the university level analysis was not as strong as the first but since it accounted for such a large portion (approx. 30%) of the test battery's discriminating power both functions were used to describe cognitive style differences among the three groups. The correlations given in Table 2 indicate these style differences when interpreted in conjunction with the relative positions of the groups in Figure 1(a).

In fact, the findings at the university level proved to be somewhat of a surprise. Function I provided a clear separation of the high (science) and low (history) groups together from the middle (engineering) group. Both the high and low groups scored well in terms of conceptual differentiation and categorising flexibility with objects (the ability to see an object as a member of a number of different groupings), and they showed a strong tendency to use analytic-descriptive concepts. On the other hand, the middle group demonstrated a very distinct preference for categorical-inferential

TABLE 2

Correlations between Original and Discriminant Variables - University II, Form VI and Form IV Levels

Variable	University II Fn I	University II Fn II	Form VI	Form IV
Conceptual Differentiation				
- Objects	.47	-.14	.48	.43
- Persons	.47	-.06	.37	.37
Categorising Flexibility				
- Objects	.46	-.26	.49	.53
- Persons	.23	.22	.27	.34
Conceptual Preference				
- Objects : % D+	.50	-.03	.31	.61
% I	-.49	.12	-.29	-.47
% Re	-.04	-.30	-.03	-.27
- Persons : % P	.26	.27	-.35	.11
% Ro	-.58	.02	.20	-.33
% S	.16	-.37	.29	.15
Originality	.23	.34	.45	.47
Ideational Fluency	.26	-.61	-.19	.09
Flexibility	.28	-.43	.57	.52
Category Width				
- Total	-.01	.09	.28	.23
- Inconsistency	-.22	.10	-.45	-.65
Latent root	0.589	0.244	0.465	1.098
Percentage of trace	70.5	29.2	98.7	99.3
χ^2 value (df)	37.5(16)	17.7(14)	42.2(15)	103.4(15)
p	<.01	NS	<.001	<.001

+ D = Descriptive; I = Inferential; Re = Relational; P = Physical; Ro = Role; S = Psychological concepts.

concepts that emphasise the use or location of objects, and for role concepts that group people in terms of possible occupations or statuses. It seems that the engineers' thinking was largely directed to practical situations as might be expected from their vocational interests. The main effect of this function, then, appears to be the discrimination of a type of analytical tendency among the science and history students from the more practical orientation of the engineers - it acts as a pure vs. applied dimension.

The three groups were also separated along a second dimension, described by Function II, but here their order did agree with the science orientation labelling (see Figure 1(a)). Since the engineering group was almost mid-way between the other two groups on this dimension, the variables listed in the lower half of Table 2 really show the major differences between the high (science) and low (history) groups. The science students showed more originality, but were less flexible and less fluent in their thinking in comparison with the history group. There was also a greater tendency on the part of the history students to group objects in terms of relational concepts (reasons for grouping that link objects together in story-like themes), and to group people on the basis of psychological concepts (i.e. in terms of inferred feelings and attitudes).

It must be stressed, however, that the differences in cognitive style described here only have their full meaning when considered as *systems* of differences operating together. It is also important to recognise that the differences are *relative*, and were obtained from the comparison of very select groups. In contrast, the school level comparisons involved groups that differed largely in general ability, and yet the cognitive style patterns obtained for the high science orientation groups showed a remarkable similarity to each other and to that of the university science group.

The school science specialists, in both Form IV and Form VI, were characterised by high conceptual differentiation, high categorising flexibility with objects, high preference for analytic-descriptive concepts, high originality, and low preference for categorical-inferential concepts. In comparison with their respective low groups they also showed high flexibility but, in fact, their scores on this variable were almost identical with that of the university science group. The three high groups also had very similar ideational fluency scores but at the two school levels this variable did not enter into the cognitive style descriptions. One variable that was important at the school levels but not among the university groups was a measure of consistency in defining category boundaries or limits; it seems highly likely that this finding was a reflection of differences in general ability among the groups of school pupils.

In summary, then, the three analyses can be taken to indicate that, for males, high science orientation (science specialisation) is associated with a specific pattern of thinking abilities. This pattern - a 'science cognitive style' - includes: high conceptual differentiation; high categorising flexibility with objects; high preference for analytic-descriptive concepts; high originality; low preference for categorical-inferential concepts; low ideational fluency; low flexibility; and low preference for psychological concepts in grouping people.

No claim is made, of course, that this cognitive style description is either complete or absolute. Many other types of thinking abilities are likely to be involved and future research will have to try to identify them. Some of the variables selected in the present study did not discriminate between the various levels of science orientation. This will have to be investigated further as also will differences in cognitive style between science specialists and comparison groups other than those used here at the university level. Of the many possibilities for future research two areas are of immediate interest: cognitive styles among female students; and cognitive styles among teachers. Are science teachers like scientists in their thinking or are they more like other teachers? Do science teachers identify with their subject matter or with the teaching profession?

Some Implications

If the findings presented here are confirmed and extended by future research, it is possible that the notion of cognitive styles may be of considerable importance as a guidance technique. The findings, even as they stand, can also be seen as having implications for classroom teaching since they identify some major differences in the thinking of high and low ability science students.

The high ability students demonstrated a style of thinking that suggests their need for a particularly challenging classroom environment. Their competence for making fine differentiations and their keen awareness of the intrinsic qualities of physical objects appear highly suited to relatively unstructured laboratory activities in which they could set about gathering their own data. If left to decide for themselves the factors that are to be observed and the order in which observations are to be made, such students would have a greater opportunity to exercise their originality and flexibility; and in finding new ways of looking at things they would increase their comprehension of the system under study. It seems equally important that these students be encouraged to apply their skills to verbal materials and to use books and journals as sources of information. In this way they could discover more for themselves and learn to rely on their teachers less. These suggestions are offered mainly with sixth-form students in mind but the present study indicates that talented science students at the fourth-form level would benefit from similar procedures provided, of course, that allowance is made for their lesser maturity.

On the other hand, the findings suggest that the science teacher faces an even more challenging task with the lower achieving students. These groups displayed markedly limited abilities on most of the thinking measures and it would be unrealistic to expect such students to formulate and test hypotheses in any detailed way by themselves, or to gather much information independently. If anything, the teacher's approach to these groups should be even more stimulating than that to more able students. Information should be constantly related to the pupils' experience, and the teacher should be continually striving to broaden the students' thinking and to help them become more aware of possible alternatives and implications. The argument being presented is that if the less able student can be helped to adopt different strategies in his thinking, his understandings of science concepts and principles might develop more readily.

The suggestions made here obviously necessitate a very open and flexible approach on the part of the teacher, and have important implications for the training of science teachers. In terms of the way students think, the science teacher should not only be an information-giver, but should also be willing to listen, and to participate in discussions. His questioning should be varied and should seek to encourage students to explain, or analyse, or compare data rather than simply requiring them to state a fact or to identify something by name. His lesson planning should be imaginative and, as always, he should be involved in a constant search for new and unusual problems of relevance to the pupils.

As emphasised at the beginning of this paper, it is essential that teachers understand more clearly how their pupils think. Much of science teaching must be concerned with the development of knowledge skills - the facts, concepts, and principles upon which the structure of science is based - but more attention needs to be given to the ways in which these skills are developed. That is, science teaching should endeavour to foster the thinking skills upon which the continued growth of the students' understandings in science depends.

REFERENCES

Barron, F. (1965). The psychology of creativity. In (Newcomb, T.M., (Ed.) *New Directions in Psychology II.* New York: Holt, Rinehart and Winston.)1-134.

Bush, M. (1969). Psychoanalysis and scientific creativity. *Journal of the American Psychoanalytic Association, 17,* 136-190.

Christenson, P.R., Merrifield, P.R., and Guilford, J.P. (1962). *Consequences: Manual for Administration, Scoring and Interpretation. (Second Edition.)* Beverly Hills, Calif.: Sheridan Supply Co.

Gardner, R.W., Jackson, D.N., and Messick, S.J. (1960). Personality organisation in cognitive controls and intellectual abilities. *Psychological Issues,* 2, No. 4 (Whole No. 8).

Kagan, J., Moss, H.A., and Sigel, I.E. (1963). Psychological significance of styles of conceptualisation. In (Wright, J.C., and Kagan, J. (Eds.). Basic cognitive processes in children. *Monograph of Society for Research in Child Development, 28,* No. 2). 73-124.

Pettigrew, T. (1958). The measurement and correlates of category width as a cognitive variable. *Journal of Personality, 26,* 532-544.

Veldman, D.J. (1967). *Fortran Programming for the Behavioral Sciences.* New York: Holt, Rinehart and Winston.

Wallach, M.A., and Kogan, N. (1965). *Modes of thinking in Young Children.* New York: Holt, Rinehart and Winston.

Wilson, R.C., Christenson, P.R., Merrifield, P.R., and Guilford, J.P. (1960). *Alternate Uses, Form A: Manual of Administration, Scoring and Interpretation. Second Preliminary Edition.*

Witkin, H.A. (1964). Origins of cognitive styles. In (Scheerer, Constance (Ed.). *Cognition: Theory, Research, Promise.* New York: Harper and Row). 172-205.

EXTERNAL FACTORS IN THE LEARNING OF GRAPHICAL SKILLS IN KINEMATICS

R.T. White

Factors which might affect the outcomes of a course of instruction can be divided into those internal to the learner, such as his knowledge, intelligence, or intrinsic motivation, and external ones such as the mode of instruction, the length of time involved in it, or the amount of assistance the instruction provides. There can be interactions between these factors. This paper is restricted to the effect of two external factors - the amount of guidance given to the learner and the sequence in which parts of the subject-matter are presented to him - on learning, retention, and transfer of graphical skills in kinematics.

Literature on the effect of amount of guidance

One of the great questions of the past decade has been the relative merits of discovery and reception learning. It seems reasonable to equate discovery methods with those in which little guidance is given to the learner, and reception with those in which a maxiumum of guidance is given. The debate on discovery and reception may be seen in full flower in the writings of Bruner (1960, 1961, 1966) and Ausubel (1963, 1968). Bruner (1961) claimed several advantages for discovery learning, which Ausubel (1968) rebutted.

The research evidence supporting the opposing theoretical positions is neither conclusive nor impressive. Table 1 is a summary of the claimed results of studies. Three stages in amount of guidance - low, medium, and high - are shown in Table 1; where a study involved only two instructional methods they have been classified as high and low guidance. Only three outcomes, initial learning, retention, and transfer, are represented in Table 1, for there appears to have been hardly any research into the effects of the methods on other outcomes.

There is a temptation to use Table 1 to draw conclusions about the relative merits of large and small amounts of guidance. This should be resisted, for several reasons:

(1) In compiling the table some studies may have been overlooked.

(2) The studies are restricted in the main to quantitative subject matter.

(3) Definitions of treatments vary from experimenter to experimenter. One man's discovery method may be another's reception. A close comparison of the studies of Craig (1956) and Kittell (1957) will illustrate this; or, on consideration of the treatments used by Kersh (1962), it is possible to make out a good argument for placing them in reverse order on a continuum from low to high guidance to that proposed by Kersh.

(4) Some of the experimenters, e.g. Hendrix (1947) and Haslerud and Meyers (1958), made claims which are not supported by their results.

(5) Many of the studies are poorly designed. They have been criticised by Olson (1965), Cronbach (1966), Ausubel (1968), Hermann (1969), and Bittinger (1968).

In the experiment described below, attempts were made to avoid these faults of vague definition of methods, poor design, and unjustified claims.

TABLE I

Conclusions of Studies in which Methods Varying in Amount of Guidance were Compared with Respect to Their Effectiveness in Promoting Initial Learning, Retention, and Transfer

Note: 1. where the study involved only two methods, they have been designated high and low amounts of guidance;
2. where the difference between methods was not significant at the .05 level, the words high, medium, and low are underlined.

Author	Ss	Subject matter	Most suitable amount of guidance for initial learning	retention	transfer
Fowler (1931)	246 6th graders	artificial and grammatical concepts	high		
Hendrix (1947)	40 11th & 12th graders and college students	number rules			low
Craig (1956)	106 college students	word principles	high	high	low
Sobel (1956)	312 9th graders	algebra	low	low	
Smith et al (1956)	80 undergraduates	reading	high		
Kittell (1957)	132 6th graders	word principles	high	medium	medium
Forgus & Schwartz (1957)	39 undergraduates	artificial alphabet		medium	medium
Kersh (1958)	60 undergraduates	numerical rules	high	low	
Haslerud & Meyers (1958)	76 undergraduates	codes	high		low
Ray	117 9th graders	micrometer	high	low	low
Gagné & Brown (1961)	33 9th & 10th graders	sums of series			medium
Gagné et al (1962)	136 7th graders	addition	high		low

Table 1 continued

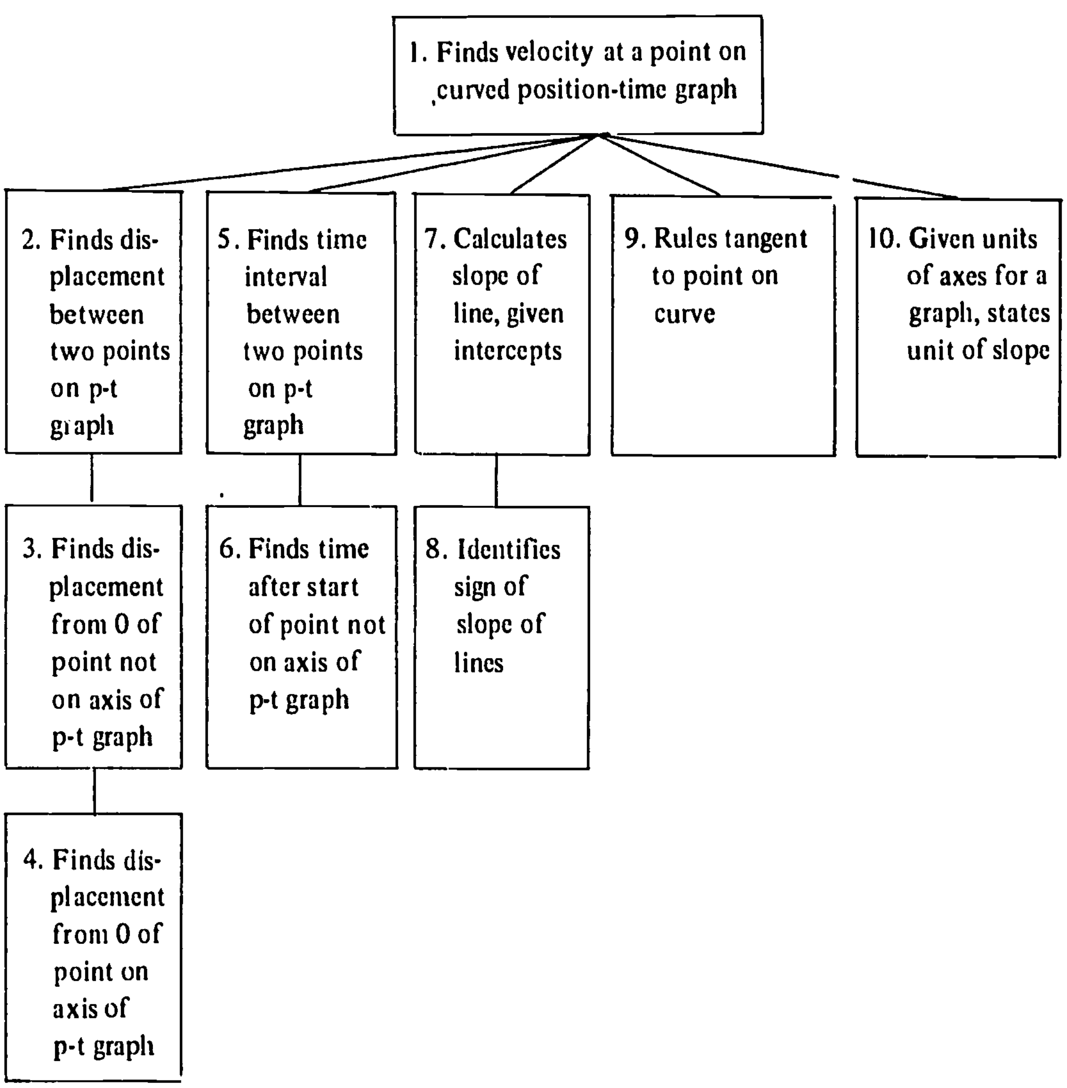
Author	Ss	Subject matter	Most suitable amount of guidance for initial learning	retention	transfer
Kersh (1962)	90 high school students	numerical rules		high	high
Wittrock (1963)	292 undergraduates	codes	high	medium	medium
Orton, McKay & Rainey (1964)	40 3rd graders and 12 mentally handicapped children	arabic and roman numbers	low	high	low
Belcastro (1966)	378 8th graders	algebra	high		
Guthrie (1967)	72 college students	cryptograms	high	high	low
Boleratz (1967)	211 9th & 10th graders	value concepts	low		low
Meconi (1967) - 1st experiment	24 8th & 9th graders	sums of series		low & medium	medium
Meconi (1967) - 2nd experiment	21 8th & 9th graders	sums of series		low	high
Werdelin (1968)	178 6th graders	multiplication rule	high	low	low
	174 8th graders	arabic alphabet	high	low	low
Worthern (1968)	432 5th & 6th graders	mathematics	high	low	low
Jamieson (1969)	80 women, ages 10 to 66	binary number	low		
Kornreich (1969)	99 undergraduates	concept - identifying strategy	medium		
Thomas & Snider (1969)	140 8th graders	inquiry skills	high		
Rowell, Simon, & Wiseman (1969)	59 undergraduates	schemata	high	high	high

Subject matter of the experiment

In an earlier study of learning of kinematics (White 1971), it was shown that the graphical skills could be represented by a learning hierarchy, such as were first investigated by Gagne (1962, 1968; Gagne and Paradise 1961). Part of this hierarchy is shown in figure 1. The symbolism of figure 1 means that skill 1 cannot be acquired unless the learner possesses, and can recall, all of skills 2, 5, 7, 9, and 10, which are known as relevant subordinate skills to skill 1. Skill 2 in turn could only have been learned if its relevant subordinate skill, skill 3, was possessed and recalled.

Figure 1

Hierarchical Arrangement of Ten Skills



Examples of questions which test the ten skills are given in figure 2.

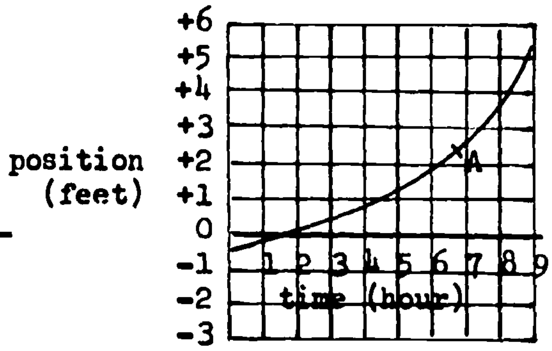
Figure 2

Questions for Skills

Skill 1

What is the velocity at A?

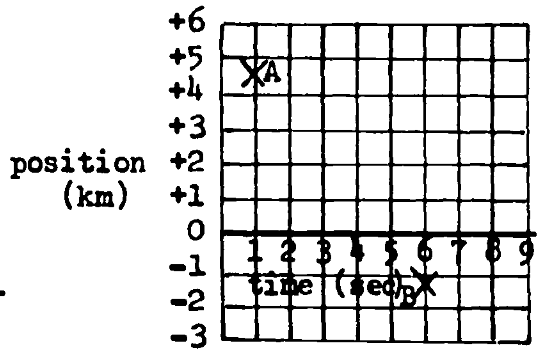
Answer: ______ ______ ______
+ or - no. unit



Skill 2

What is the displacement in going from A to B?

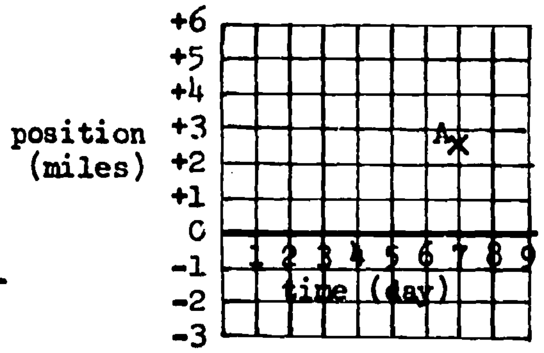
Answer: ______ ______ ______
+ or - no. unit



Skill 3

What is A's displacement from 0?

Answer: ______ ______ ______
+ or - no. unit



Skill 4

What is A's displacement from 0?

Answer: ______ ______ ______
+ or - no. unit

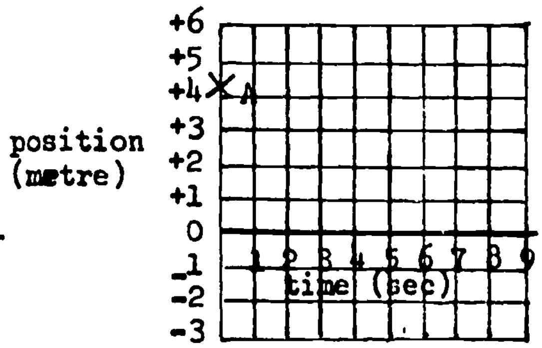
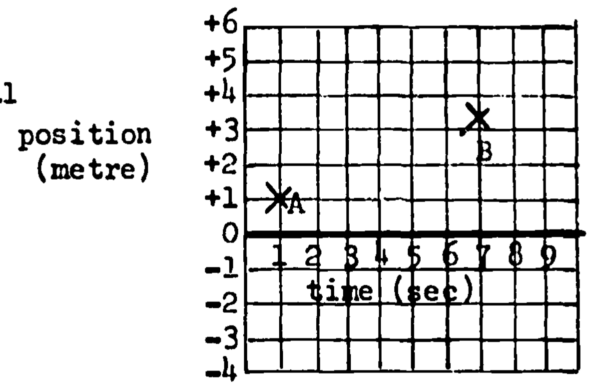


Figure 2 cont.

Skill 5

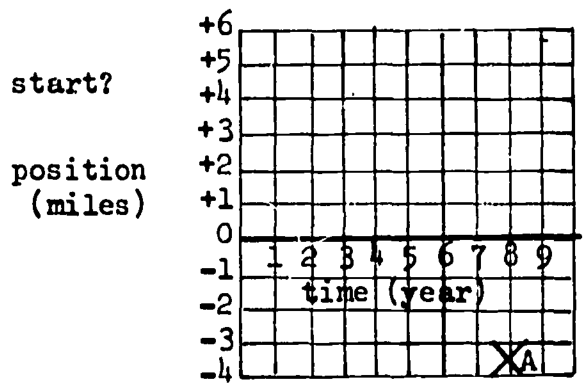
What is the time interval between A & B?



Answer: ____ ____
no. unit

Skill 6

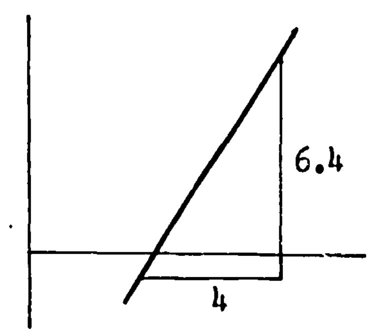
What time is A after the start?



Answer: ____ ____
no. unit

Skill 7

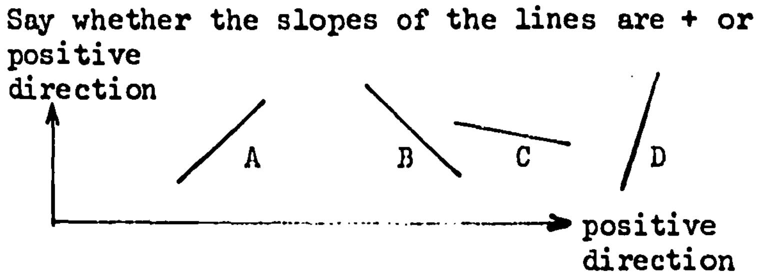
What is the slope of this line?



Answer: ____ ____
+ or - no.

Skill 8

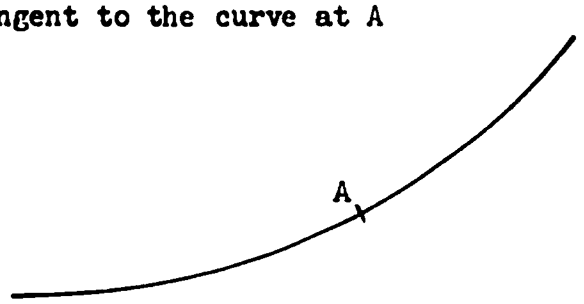
Say whether the slopes of the lines are + or -.



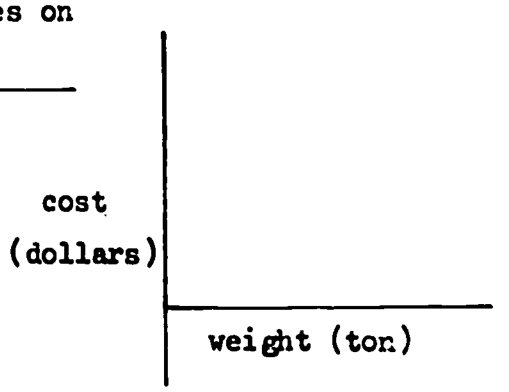
line	+ or -
A	
B	
C	
D	

Figure 2 cont.

Skill 9
Rule a tangent to the curve at A



Skill 10
The unit of the slope of lines on this graph is ____________



The three factors in the experiment

The main purpose of the study was to investigate the effects of four methods, which varied in the amount of guidance they were designed to provide, on learning, retention, and transfer of the ten skills. The four methods, from least to most guidance, were:

Method A. The learner was given a question testing the skills, and when he made a response he was told whether his answer was correct, and, if not, the right answer.

Method B. The same as A, with the addition of a sheet of paper containing a reminder of relevant subordinate skills. This paper was given to the learner at the same time as the first question, and was left in his possession.

Method C. The same as B, with the addition of instructions on the sheet of paper on how to answer the question.

Method D. The same as C, with the addition of a worked example on the sheet of paper.

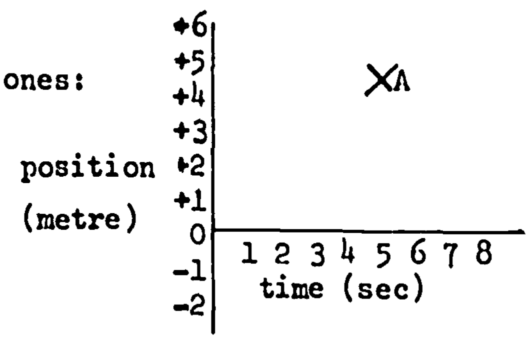
Figure 3 shows the instruction sheets for methods B, C, and D for skill 2.

Figure 3

Three Forms of Guidance for Skill 2

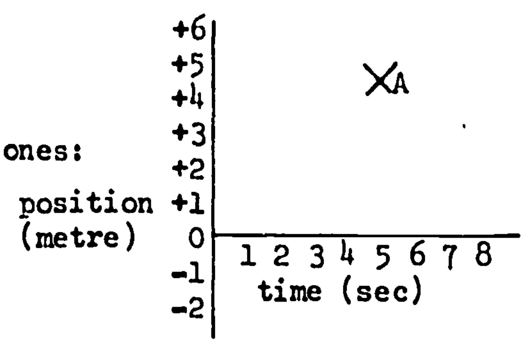
Reminder only

Remember how you did these ones:



Reminder & instructions

Remember how you did these ones:



In the new ones, find the points A & B are level with, and see how far apart these points are

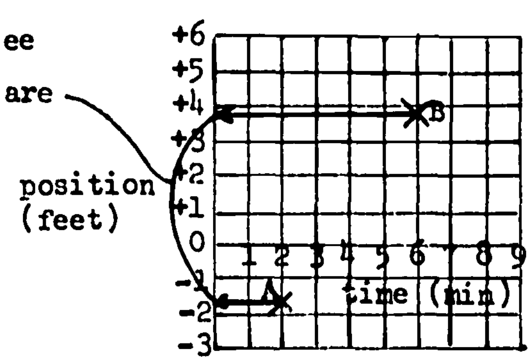
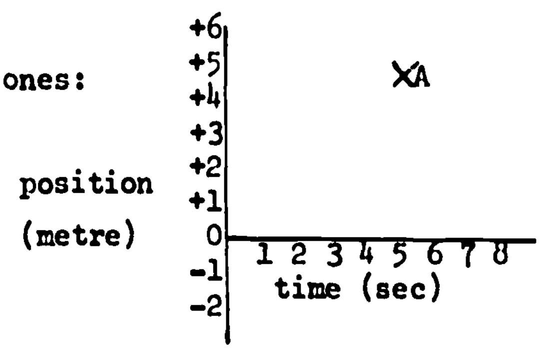


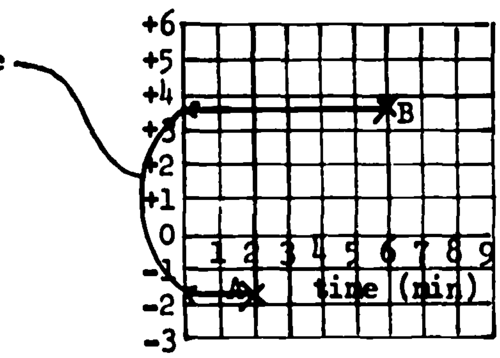
Figure 3 cont.

Reminder, instructions, & example

Remember how you did these ones:

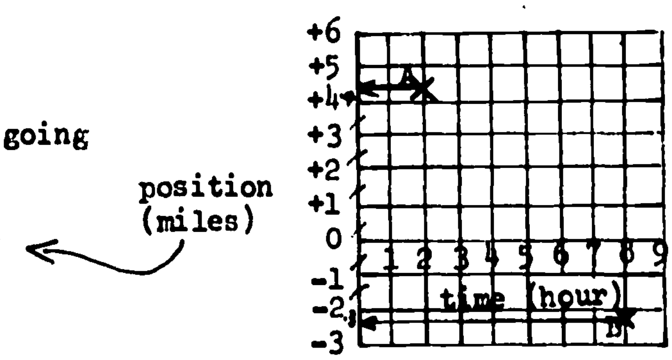


In the new ones, find the points A & B are level with, and see how far apart these points are



Here is an example to help you:

The displacement in going from A to B is



–	6·7	miles
+ or –	no.	unit

The experiment was actually a three-factor one. One factor, with four levels, was the method used to teach the relevant subordinate skill, which had four levels also. This was included in the experiment as an investigation of transfer. Transfer in this sense is what Gagné (1965, p.231) has distinguished as vertical transfer. It is connected with one of the claims made by Bruner (1961) for discovery learning, that knowledge which is self-acquired will be more readily available for use in new learning. This leads to the prediction that subjects who learn skill 3, for example, by method A will learn skill 2 with fewer errors than other subjects, all other factors being equal.

The third factor, with two levels, was the sequence in which the ten skills were attempted. Both types of sequence had to be in accord with the requirements of the learning hierarchy, that skill 4, for example, had to be learned before skill 3 was presented, but differed in whether the skills in each of the five ladders in the hierarchy (4-3-2 is one ladder; 6-5, 8-7, 9, and 10 are the others) were taught in succession or not. These two sequence types were called ladder and cross respectively. Examples of ladder sequences are:

10-9-8-7-6-5-4-3-2-1
6-5-10-4-3-2-8-7-9-1

Examples of cross sequences are:

4-6-3-10-8-5-2-7-9-1
9-6-5-4-8-10-3-7-2-1

Administration of the experiment

One hundred and twenty-eight subjects were fitted into the 4 x 4 x 2 design. Their allocation to methods was random, with a separate allocation to each skill, subject to the constraint that there were to be equal numbers of subjects in each of the thirty-two cells for each skill. For skills 4, 6, 8, 9, and 10 the design regressed to 4 x 2, as these skills are at the bottoms of the ladders and, although the original hierarchy (White 1971) did include skills below them, in this experiment they had no subordinate skill that was taught by the four methods A, B, C, or D.

The 128 subjects were drawn randomly from the form two rolls of two Melbourne metropolitan high schools, which also were chosen randomly. The subjects were taken through the experiment in groups of four or five at a time, by the one experimenter. They were taught each skill in turn until they reached a criterion performance. Each subject was given the appropriate instruction sheet for method A, B, C, or D, together with a question testing the skill. After he had attempted the question, the experimenter told him whether he was correct or not, and if he was wrong, what the right answer was. He was then given the next question for the same skill. This was continued until the learner reached the criterion of three correct answers in succession. He was then given the instruction sheet and first question for the next skill. The dependent variable in the experiment was the number of errors each learner made before reaching the criterion.

Three subjects were replaced by emergencies, also randomly chosen, because one could not read, one failed to learn how to draw a tangent (skill 9) after ten attempts with method C, and one failed to learn skill 4 after sixteen attempts with method A. All other subjects reached the criterion for skills 2 to 10. Skill 1 was difficult, and if a subject had not reached the criterion for it after two hours total instruction the experiment was terminated.

Retention test

The subjects were not told there was to be a retention test, but eight weeks later they were reassembled for one. They were given a set of questions similar to those in figure 2, one for each skill, except skill 1, then another set in different order, then a third and a fourth set. Four sets were used in case answering the question for one element triggered recall of another, which may have been learned by a different method. One-quarter of the subjects began with set 1, one-quarter with set 2, and so on. Some improvement was noted from the first-answered set to the third, but then the performance became stable. Skill 1 was not included in the test because different numbers of subjects from the four methods groups, and also from the four previous methods groups for each relevant subordinate element and the two sequence groups, had learned this skill. Therefore the groups could not be expected to be equivalent with respect to memory of the skill. For the other nine skills this difficulty did not exist, since all subjects in all groups did learn them, so that the groups had equal chances, apart from the specific effects on memory of the treatments, of retaining the skills.

Results

The numbers of errors made in the learning experiment were highly positively skewed, so the non-parametric Kruskal-Wallis test was used in preference to parametric analysis of variance. The extension of the Kruskal-Wallis test to multi-factor experiments described by Bradley (1968) was applied.

Table 2 shows mean numbers of errors made by subjects under the four methods, for skills 2 to 10. The H value is that obtained in the Kruskal-Wallis analysis, and probabilities, p, of obtaining H values as large or larger than these, given that the methods do not differ in effect, are also shown.

Table 3 shows the mean numbers of errors made in learning a skill by subjects who learned the subordinate skill by each of the methods.

Table 4 shows the mean numbers of errors made for the two sequence types.

TABLE 2

Effects of Methods on Initial Learning

Skill	Mean nos. of errors for method A	B	C	D	H	p
2	6.28	5.31	2.69	2.06	10.73	<.01
3	.78	.72	.56	.34	4.75	>.05
4	.81	1.08	1.16	.38	13.99	<.01
5	1.38	1.47	.63	.38	10.64	<.025
6	1.41	1.34	1.19	1.09	2.75	>.05
7	3.72	2.69	1.13	.44	13.79	<.01
8	1.44	1.84	.41	.31	22.99	<.01
9	1.03	.16	.03	.03	47.93	<.01
10	1.19	1.25	1.34	.09	39.93	<.01

TABLE 3

Effects of Methods on Vertical Transfer

Skill	Subordinate Skill	Mean nos. of errors for skill subordinate skill method A	B	C	D	H	p
2	3	4.22	3.91	4.09	4.13	1.23	>.05
3	4	.47	.59	.69	.65	1.36	>.05
5	6	.81	.53	1.00	1.50	9.22	<.05
7	8	2.09	1.72	1.66	2.50	3.16	>.05

TABLE 4

Effects of Sequences on Initial Learning

Skill	Mean nos. of errors for sequence Ladder	Cross	H	p
2	3.61	4.56	.08	>.05
3	.53	.67	1.33	>.05
4	.95	.77	.35	>.05
5	.80	1.13	1.69	>.05
6	1.27	1.25	0	>.05
7	2.09	1.89	.08	>.05
8	.88	1.13	1.77	>.05
9	.34	.28	.89	>.05
10	.97	.97	.06	>.05

Table 5 shows H and p values for interactions between the three factors.

Table 6 shows the number of subjects in each group for skill 1 who reached the criterion, together with X^2 and p values.

Table 7 shows how many subjects in each group remembered the skill on the fourth-answered set of questions in the retention test. Similar results were obtained for the other three sets. The probabilities of these observed distributions given that the factors have no effect on retention are also shown.

TABLE 5

Interaction Effects

1. Method x Method for subordinate element

Skill	H	p
2	9.76	>.05
3	7.59	>.05
5	16.08	>.05
7	8.53	>.05

2. Method x Sequence

Skill	H	p
2	3.01	>.05
3	5.33	>.05
4	1.90	>.05
5	2.97	>.05
6	2.36	>.05
7	1.27	>.05
8	4.62	>.05
9	1.67	>.05
10	3.99	>.05

TABLE 5 cont.

3. Method for subordinate element x Sequence

Skill	H	p
2	2.74	>.05
3	1.91	>.05
5	5.73	>.05
7	.89	>.05

4. Method x Method for subordinate element x Sequence

Skill	H	p
2	20.56	<.025
3	5.74	>.05
5	9.71	>.05
7	10.07	>.05

TABLE 6

Results for Skill 1

1. Effect of method on initial learning

	Method A	Method B	Method C	Method D
No. of subjects to reach criterion	3	6	19	21

$\chi^2 = 20.14$; $p < .01$

TABLE 6 cont.

2. Effect of method on vertical transfer

Subordinate Skill	No. of subjects to reach criterion for skill 1, for subordinate skill method				χ^2	p
	A	B	C	D		
2	10	13	13	13	.55	>.05
5	12	14	11	12	.38	>.05
7	9	14	13	13	1.20	>.05
9	14	11	13	11	.55	>.05
10	15	12	11	11	.88	>.05

3. Effect of sequence on initial learning

	Sequence	
	Ladder	Cross
No. of subjects to reach criterion	28	21

$\chi^2 = .73$; $p > .05$

TABLE 7

Effects of Methods and Sequences on Retention

1. Effect of methods

Skill	Fraction of subjects to remember Method A	B	C	D	χ^2	p
2	4/27	5/27	5/27	2/27	.30	>.05*
3	12/28	10/27	9/26	14/27	1.88	>.05
4	15/28	14/23	11/30	14/27	1.57	>.05
5	10/28	11/29	12/27	6/24	2.11	>.05
6	8/27	13/27	14/25	6/29	9.09	<.05
7	6/27	5/27	2/29	2/25	3.78	>.05*
8	17/29	16/27	14/24	19/28	.72	>.05
9	21/24	20/25	24/29	26/30	.02	>.05*
10	4/28	5/27	9/27	6/26	3.15	>.05

*In these three cases columns A and B, and C and D, were amalgamated because of small expected values.

2. Effect of sequences

Skill	Fraction of subjects to remember Sequence Ladder	Cross	χ^2	p
2	4/54	12/54	4.70	>.05
3	19/54	25/54	1.38	>.05
4	25/54	29/54	.59	>.05
5	17/54	19/54	.08	>.05
6	20/54	22/54	.16	>.05
7	8/54	7/54	.08	>.05
8	11/54	13/54	.21	>.05
9	45/54	46/54	.07	>.05
10	11/54	13/54	.21	>.05

Conclusions

It is clear from Table 2 that fewer errors are made when greater guidance is given. Number of errors is probably closely related to the time taken to reach the criterion, and therefore for efficient instruction learners should be given as much assistance as possible. Table 6 supports this.

The other tables show that the other outcomes, retention and transfer, are not affected by which method is used. They also show that sequence is a factor of little or no effect on any of the three outcomes.

Limitations

These conclusions must be tempered with warnings on the limitations of the experiment. The instructional period was only two hours, and over a longer time it is possible that motivational differences between high and low guidance methods will become more important, which might then cause differences to appear in the effects of the methods on retention and transfer. Differences might also appear if one method is used throughout for a learner, instead of the persistent flitting from one method to another which was used in the experiment. Different modes of instruction, such as live teachers, films, or programmed machines might lead to a different conclusion from that obtained in the experiment, where a programmed book was used. Sequence may have been found to be an effective factor if longer ladders had been present in the subject matter.

The general recommendation which follows from the experiment is that teachers of quantitative skills should give their pupils as much assistance as possible. This assistance could include reminders of relevant subordinate elements, detailed instructions, and worked examples. Further research into exploring the limitations of this recommendation is advocated.

The final limitation is that only three outcomes were measured. Others, such as those listed by Cronbach (1966) which include development of long- and short-term interest in the subject matter and of general problem-solving skills, were omitted because they were unlikely to be affected by such a short teaching treatment.

REFERENCES

Ausubel, D.P. (1963). *The Psychology of Meaningful Verbal Learning,* New York, Grune & Stratton.

Ausubel, D.P. (1968). *Educational Psychology: a Cognitive View,* New York, Holt, Rinehart & Winston.

Belcastro, F.P. (1966). "Relative effectiveness of the inductive and deductive methods of programming algebra", *Journal of Experimental Education, 34* (3), 77-82.

Bittinger, M.L. (1968). "A review of discovery". *Mathematics Teacher, 61,* 140-146.

Bolerat‹, J.M. (1967). "Learning by discovery: an experimental study to measure its effectiveness fnr teaching value concepts", *Journal of Experimental Education, 36* (2), 13-21.

Bradley, J.V. (1968). *Distribution-Free Statistical Tests,* New Jersey, Prentice-Hall.

Bruner, J.S. (1960). *The Process of Education,* Cambridge, Harvard University Press.

Bruner, J.S. (1961). "The act of discovery", *Harvard Educational Review, 31,* 21-32.

Bruner, J.S. (1966). *Toward a Theory of Instruction,* Cambridge, Harvard University Press.

Craig, R.C. (1956). "Directed versus independent discovery of established relations", *Journal of Educational Psychology,* 47, 223-234.

Cronbach, L.J. (1966). "The logic of experiments on discovery", Chapter 5 of *Learning by Discovery: a Critical Appraisal,* eds. L.S. Shulman and E.R. Keislar, Chicago, Rand-McNally.

Forgus, R.H. & Schwartz, R.J. (1957). "Efficient retention and transfer as affected by learning methnd", *Journal of Psychology,* 43, 135-139.

Fnwler, H.L. (1931). *Induction or Deduction? An Experimental Investigation in the Psychology of Teaching,* Melbourne, Melbourne University Press.

Gagné, R.M. (1962). "The acquisition of knowledge", *Psychological Review, 69,* 355-365.

Gagné, R.M. (1965). *The Conditions of Learning,* New Ynrk, Holt, Rinehart, and Winston.

Gagné, R.M. (1968). "Learning hierarchies", *Educational Psychologist, 6,* 1-9.

Gagné, R.M., and Brown, L.T. (1961). "Some factors in the programming of conceptual learning", *Journal of Experimental Psychology, 62,* 313-321.

Gagné, R.M., and Paradise, N.E. (1961). "Abilities and learning sets in knowledge acquisition", *Psychological Monographs, 75:* 14 (whole number 518).

Gagné, R.M. , et al (1962). "Factors in acquiring knowledge of a mathematical task", *Psychological Monographs, 76:* 7 (whole number 526).

Guthrie, J.T. (1967). "Expository instruction versus a discovery method", *Journal of Educational Psychology, 58,* 45-49.

Haslerud, G.M., and Meyers, S. (1958). "The transfer value of given and individually derived principles", *Journal of Educational Psychology, 49,* 293-298.

Hendrix, G. (1947). "A new clue to transfer of training", *Elementary School Journal, 48,* 197-208.

Hermann, G. (1969). "Learning by discovery: a critical review nf studies", *Journal of Experimental Education, 38,* 58-72.

Jamiesnn, G.H. (1969). "Learning by programmed and guided discnvery methods at different age levels", *Programmed Learning and Educational Technology, 6,* 26-30.

Kersh, B.Y. (1958). "The adequacy of 'meaning' as an explanation for the superiority of learning by independent discovery", *Journal of Educational Psychology, 49,* 282-292.

Kersh, B.Y. (1962). "The motivating effect of learning by directed discovery", *Journal of Educational Psychology, 53,* 65-71.

Kittell, J.E. (1957). "An experimental study of the effect of external direction during learning on transfer and retention of principles", *Journal of Educational Psychology, 48,* 391-405.

Kornreich, L.B. (1969). "Discovery versus programmed instruction in teaching a strategy for solving concept-identification problems", *Journal of Educational Psychology, 60,* 384-388.

Meconi, L.J. (1967). "Concept learning and retention in mathematics", *Journal of Experimental Education, 36* (1), 51-57.

Olson, D.R. (1965). "Note on Haslerud and Meyers' transfer-of-principles experiment", *Journal of Educational Psychology, 56,* 107-108.

Orton, K.D. McKay, B.E., and Rainey, D. (1964). "The effect of method of instruction on retention and transfer for different levels of ability", *School Review,* 72, 451-461.

Ray, W.E. (1961). "Pupil discovery versus direct instruction", *Journal of Experimental Education,* 29 (3), 271-280.

Rowell, J.A., Simon, J., and Wiseman, R. (1969). "Verbal reception, guided discovery and the learning of schemata", *British Journal of Educational Psychology, 39,* 233-244.

Smith, D.E.P., et al (1956). "Reading improvement as a function of student personality and teaching method", *Journal of Educational Psychology, 47,* 47-59.

Sobel, M.A. (1956). "Concept learning in algebra", *Mathematics Teacher, 49,* 425-430.

Thomas, B., and Snider, B. (1969). "The effects of instructional method upon the acquisition of inquiry skills", *Journal of Research in Science Teaching, 6,* 377-386.

Werdelin, I. (1968). "The value of external direction and individual discovery in learning situations", *Scandinavian Journal of Psychology, 9,* 241-251.

White, R.T. (1971). "Learning graphical skills in kinematics", unpublished Ph.D. thesis, Monash University.

Wittrock, M.C. (1963). "Verbal stimuli in concept formation", *Journal of Educational Psychology, 54,* 183-190.

Worthern, B.R. (1968). "Discovery and expository task presentation in elementary mathematics", *Journal of Educational Psychology, 59,* monograph supplement 1-13.

THE RELATIONSHIPS AMONG SOME COMPONENTS OF STUDENT-CENTRED INQUIRY METHODS AND STUDENTS' PERCEPTIONS OF THE OUTCOMES OF BIOLOGY TEACHING *

Effie D. Best

* This paper is based on work done at The Ohio State University in partial fulfilment of requirements for the Ph.D. Particular acknowledgement is given to the author's advisor, Dr Robert W. Howe, for his encouragement and guidance.

Introduction

Many science educators advocate "student-centred inquiry" methods in science teaching. These methods are sometimes contrasted with "authoritarian" methods (Carter 1967) implying that teachers can be aligned along a single continuum from the "traditional" teacher who emphasizes observational or confirmatory laboratory work in a teacher-centred classroom to the "inquiry" teacher who emphasizes investigative laboratory work in a student-centred classroom. A similar assumption seems to underlie instruments which have been developed to assess the extent to which classrooms conform to the student-centred inquiry model (Kockendorfer 1966).

However, it can be questioned whether there is a single continuum of science teaching methods from "traditional" to "inquiry". There seems to be no logical necessity for observational laboratory work to be teacher-centred, or for many kinds of investigative laboratory work to be student-centred. Hence one of the aims of the study reported here was to identify and measure *components* of "student-centred inquiry" methods as used by biology teachers in the field, and to investigate relationships between these components. A second aim was to investigate relationships between those components for which measures were developed and the learning outcomes as perceived by students.

Sampling Procedures

Thirty three biology teachers in three urban school districts in Ohio, USA, provided data for the study. The teachers were not randomly selected; the selection procedures were biased to include a relatively high proportion of teachers who emphasized laboratory work. Data were collected from one class of each teacher, chosen by the teacher. The sample included ninth grade and tenth grade classes. Preliminary analyses of the data indicated that there were insufficient differences between school districts or grade levels to necessitate separating these for further analyses.

Three instruments were administered to the students by the teacher after they had completed an activity during which the class was "using laboratory methods to investigate a problem". The three instruments, a *Biology Activity Report*, an *Attitude Inventory*, and a *Biology Teaching Inventory*, were sent to the teacher in a randomized stack to ensure that a random third of each class responded to each inventory. Class means were computed for all variables and used for correlational analyses.

Components of "Student-centred Inquiry"

The following criteria were identified as appropriate for judging the extent to which a classroom is inquiry-oriented and student-centred:

the extent to which problem solving is emphasized.
the quality of teacher-pupil relationships.
the extent to which students participate actively in classroom activities.
the extent to which students made decisions about classroom goals and procedures.

Measurement of Components

The *Biology Activity Report* was designed to give measures of the problem orientation of the activity students had just completed, and of students' perceptions of the origins of decisions made during the activity. Factor analysis of student responses was used to modify *a priori* scales, giving the following scales:

I scale, comprising 15 items which gave a measure of the problem orientation of the activity (Table 1).
Student Procedural Decision (SPD) scale, scored from 10 items to give a measure of the number of decisions about laboratory procedures perceived by students as made by themselves relative to the number perceived as made by the teacher (Table 2).

TABLE 1

Example Items from the I Scale

Did the activity raise new questions which were discussed in class?	(Keyed "yes")
Did you consider how these new questions could be investigated?	(Keyed "yes")
After the activity, did you discuss ways in which the procedures could have been improved?	(Keyed "yes")
The class did not consider the accuracy of the data.	(Keyed "false")
There were some unexpected results, but we ignored them.	(Keyed "false")

Written Instructions (WI) scale, scored from the same 10 items to give a measure of the extent to which students perceived decisions as dictated by written instructions (Table 2).
Team scale, comprising 10 items, giving a measure of the extent to which students perceived decisions as made by individuals or teams and not by the whole class (Table 2).
Closure scale, comprising nine items, giving a measure of the extent to which all students in the class agreed on results and conclusions and of the extent to which the teacher was perceived as evaluating results and interpretations (Table 3).
Teacher Evaluation, three items from the Closure scale related specifically to the teacher's evaluation of results and interpretations (Table 3).

TABLE 2

SPD Scale, WI Scale and Team Scale Items

The following parts of the activity were listed twice:

General problem investigated
Parts of the problems studied, or hypotheses tested
General outline of procedures
Kind of data collected (e.g., *what* was measured or observed)
Exact way data were collected (*how* and *when* you measured or observed)
Amount of data collected (e.g., how many set-ups, how many observations)
Method of recording data while making observations
Method of organizing data (e.g., in graphs or tables)
Precautions taken to make sure data were as accurate as possible
Method of controlling the experiment

For *SPD scale* and *WI scale* students marked the above according to the following Key. SPD scale scoring is shown in brackets:

1 thought up by students and decided by students (3)
2 thought up by students and decided by the teacher (2)
3 given by the teacher (1)
4 given by written instructions (2)
5 was not part of the activity (deleted from computation)

For *Team scale*, the key was:

1 decided by the class (0)
2 decided by you or your team (1)
3 neither of the above apply (0)

TABLE 3

Example Items from the Closure Scale

Did those students or those teams who were using the name procedures all get the same results? (Keyed "yes")

* Did the teacher say whether the results were correct? (Keyed "yes")

* Did the teacher say whether the explanations of the results were correct? (Keyed "yes")

Did all members of the class agree on the conclusions at the end of the activity? (Keyed "yes")

* Did the teacher say whether the conclusions were correct? (Keyed "yes")

---

* These items were also scored separately as Teacher Evaluation.

TABLE 4

Example Item from TPRS

Is your biology teacher more interested in students as people or in teaching them the subject?

A Interested in each student as a person; understands different students' needs, abilities, and interests; aware of students' feelings and problems; tries to help each student wherever the student needs help.

B Is aware of different needs and problems of students, but does little to help them except in learning the subject; is quite willing to accept different standards of work from students with different abilities.

C Seems to be aware of different needs and problems of students, but believes that the teacher is only responsible for students' understanding of the subject; knows that students have different abilities but doesn't do much about those differences.

D Doesn't realize that students have problems except with the subject; takes little notice of differences in students' abilities, but does try to make lessons interesting.

E Thinks in terms of learning subject matter only; expects every student to do equally good work; students are expected to learn everything even if it makes no sense to them.

The *Attitude Inventory* included a six-item Teacher-Pupil Relationships Scale (TPRS), based on an instrument used by Howe (1964), which gave a measure of students' perceptions of their teacher as student- or subject-oriented (Table 4).

No precise measure of student involvement was used in this study.

Reliability estimates were calculated for the *Biology Activity Report* scales from item intercorrelations (Guilford, 1965, p. 463), and for the Teacher-Pupil Relationships Scale using analysis of variance (Hoyt, 1941) with the results shown in Table 5. Further data on the reliability and validity of the instruments used are given elsewhere (Best, 1970).

TABLE 5

Reliability Estimates of Scale

Scale	No. of items	Reliability
SPD scale	10	.80
Team scale	10	.82
I scale	15	.74
WI scale	10	.81
Closure scale	9	.69
TPRS	6	.85

Relationships between Components

Intercorrelations of class means on the *Biology Activity Report* scales and their correlations with class means on the Teacher-Pupil Relationships Scale are shown in Table 6.

TABLE 6

Intercorrelations of Biology Activity Report Scales and Correlations with TPRS

	SPD	Team	I scale	WI scale	Closure	T. Eval.
TPRS	.026	.221	.268	.307*	.171	.198
SPD		.392**	.017	.538***	-.316*	-.211
Team			.189	-.039	.157	.155
I scale				.173	.458***	.487***
WI scale					-.231	-.195
Closure						.900***

* p<.10 with 31 d.f.
** p<.05 with 31 d.f.
*** p<.01 with 31 d.f.

It can be seen from Table 6 that in this sample the problem orientation of the activity as measured by the I scale was not related to the extent to which students made decisions as measured by the SPD scale or Team scale. The positive correlation between I scale and Teacher Evaluation does not necessarily mean that a higher *proportion* of evaluative decisions was made by the teacher during more problem-oriented activities, because it seems likely that the total number of evaluative decisions made would be greater during investigative than observational activities. However it does suggest that few teachers deliberately refrained from evaluating results and conclusions during investigative activities.

Similarly, it is clear from the correlations of the *Biology Activity Report* scales with the Teacher-Pupil Relationships Scale that there was no significant relationship between the students' perception of the teacher as student - rather than subject-oriented and the extent to which they perceived themselves to have made decisions during the laboratory activity.

The strength of relationships found in these data cannot of course be generalized beyond the sample studied. However the lack of significant correlations between student decision making and either the students' perceptions of teacher-pupil relationships or the problem orientation of the activity means that teachers who emphasize problem solving or are perceived as student-oriented do not automatically encourage students to make decisions. It appears that these are sufficiently independent components of student-centred inquiry to be worthy of separate study.

Relationships between Measured Components and Outcome and Attitude Measures

Outcome and attitude measures were derived from the free response *Biology Teaching Inventory* and from two forms of the Silance *Scale for Measuring Attitude Toward Any School Subject* (Shaw and Wright, 1967) which were incorporated into the *Attitude Inventory*.

Outcome Measures

The free response Biology Teaching Inventory included questions about what students considered to be the most important reasons for learning biology and about the main ways in which they felt they had profited from biology classes during the year. Answers to these questions were considered to indicate the important outcomes of learning biology as perceived by the students.

Students' outcome statements were classified using a category system based on one developed by Hough and Duncan (1970) to classify instructional objectives. The Hough-Duncan system has two dimensions: *domain* (cognitive, affective and psychomotor) and *level* (a basic unit in each domain - knowledge, feeling state, or motor skill - and convergent, divergent or evaluative applications of the basic unit). A third dimension, *content*, was added to the category system for this study, divided into subject matter, process, and psycho-social. Table 7 shows the main domain and content dimensions. The level dimension was modified slightly for this study, and a fourth domain, "perceptual" was added, but these categories will not be discussed in this paper.

TABLE 7

Domains and Content Areas of the Category System

		Content Area		
Domain	Basic Unit	Subject Matter	Process	Psycho-social
Cognitive	knowledge	of biological facts and generalizations	of learning and problem-solving processes	of self or other people
Affective	feeling states	about biology or about natural phenomena	about learning and problem-solving processes	about self or other people
Psychomotor	motor skills	related to manipulation of biological materials or apparatus		related to social inter-action

Outcome statements made by students were coded for *domain, content,* and *level.* Statements such as "I didn't profit" were coded "No Value", and statements such as "to get credit" or "I needed it to graduate" were coded "Formal". The number of statements made in each category was counted for each class, then divided by the number of students responding from the class, then divided by the number of students responding from the class to give class mean scores for the categories. The structure of the category system made it possible to sum categories, for example to consider the sum of affective outcomes, or the sum of cognitive process outcomes.

Coder stability was checked by re-coding all statements from eight classes (randomly selected) and calculating percentage agreement, giving .91 for domain, .90 for content and .84 for level.

In interpreting the results of analyses of data, it should be remembered that the instrument used measured *students' perceptions* of what they had learned. This technique appeared to give a valid measure of affective outcomes difficult to measure by other techniques. For example when a student stated "I learned not to be afraid of animals, like worms and insects and snakes" (classified as affective subject matter) one could be fairly sure that this was an affective outcome of the course. On the other hand the technique allowed no assessment of quality: a student who reported having learned to solve problems may have been a rather poor problem solver. But perhaps the fact that he believed that he could solve problems as a result of learning biology is of educational significance.

Attitude Measures

The students' general attitudes to biology were assessed using seventeen items from Remmers' short form of the Silance *Scale for Measuring Attitude Toward Any School Subject, Form A* with "this subject" replaced by "biology". General attitudes to biology laboratory work were assessed by twelve items from the same scale with "this subject" replaced by "biology laboratory work". The number of "No Value" responses given on the *Biology Teaching Inventory* was also considered to indicate class attitudes to biology. The correlation between class means for the two Silance scales were -.738 ("No Value" and positive attitudes to biology) and -.741 ("No Value" and positive attitudes to biology laboratory work).

Correlations among Teaching Variables and Outcome and Attitude Measures

Correlational and multiple regression analyses were performed to investigate relationships among the variables measured. In the regression analyses, outcome measu s were treated as dependent variables and *Biology Activity Report* Scales and the Teacher-Pupil Relationships Scale as independent variables. Attitude measures were treated as dependent variables in some analyses and an independent variables in others.

TABLE 8

Correlations between Teaching Variables and Attitudes

Attitudes	TPRS	SPD	Team	I scale	WI scale	Closure
Attitude to Biology	.760***	.057	.388**	.380**	.181	.394**
Attitude to Biol Lab	.567***	.063	.350**	.277	.100	.299*
No Value	-.610***	-.129	-.345**	-.331*	-.193	-.262

* $p < .10$ with 31 d.f.

** $p < .05$ with 31 d.f.

*** $p < .01$ with 31 d.f.

TABLE 9

Correlations of Outcome Measures with Teaching Variables and Attitudes

Outcomes	SPD	Team	I scale	VI scale	Closure	TPRS	Att. to biol.	Att. to lab
Cognitive (sum)	.327*	.332**	.275	.241	-.029	.208	.338	.518***
Cognitive Subject Matter	.039	.243	.137	-.048	.007	.047	.176	.446***
Cognitive Processes	.513***	.393**	.306*	.467***	-.056	.268	.346**	.242
Affective (sum)	.232	.187	.145	.434**	.241	.629***	.573***	.320*
Affective Subj. Matter	.187	.021	-.014	.473***	.125	.480***	.456***	.305*
Affective Psycho-social	.224	.252	.348**	.092	.068	.405**	.257	.048
Psychomotor (sum)	-.242	.276	.258	-.099	.251	.372**	.322*	.170

* $p < .10$ with 31 d.f.

** $p < .05$ with 31 d.f.

*** $p < .01$ with 31 d.f.

· Two interesting patterns which emerged from these analyses were the relationships between teacher-pupil relationships and affective measures and those between student decision making and cognitive process outcomes.

All of the affective variables measured - attitudes to biology and to biology laboratory work as measured by the Silance scales, the number of "No Value" responses, and the number of affective responses given in all domains and content areas - were related to students' perceptions of teacher-pupil relationships. The regression analyses indicated that these relationships were relatively independent of any relationships between the *Biology Activity Report* scales and the outcome measures.

Cognitive process outcomes given by students included outcomes related both to problem solving processes and to learning processes appropriate to formal learning situations. The regression analyses suggested that the two student decision variables, SPD scale and Team scale, were correlated generally with process outcome, and that I scale and WI scale were correlated more specifically with problem solving processes.

It is noteworthy that none of the teaching variables measured in this study was significantly related to cognitive subject matter outcomes. It seems to be a fairly general finding that different teaching approaches rarely result in significant differences in basic subject matter knowledge. The feeling of many teachers and educators that differing approaches produce different gains in other outcome areas gains some support from this study.

Conclusions

The results of this study suggest that an analytic approach to classroom research is worth pursuing. It was found possible to identify and measure components of "student-centred inquiry" teaching which were not highly intercorrelated. It was also found that these components were differentially related to student outcomes. It would be desirable to establish whether the same relationships hold in different educational situations or when different measures are used. In particular, the relationships between teacher-pupil relationships and affective outcomes and between student decision making and cognitive process outcomes seem worthy of further investigation.

REFERENCES

Best, Effie D. (1970). *An Exploratory Study of the Correlates of Student Decision Making in the Secondary School Biology Laboratory*, Doctoral Dissertation, The Ohio State University.

Carter, Jack L. (1967). "The Authoritarian vs. The Inquiry Approach", *School Science and Mathematics, 67*, 868-9.

Guilford, J.P. (1965). *Fundamental Statistics in Psychology and Education.* 4th ed. New York, McGraw-Hill Book Company.

Hough, John B. and Duncan, James K. (1970). *Teaching: Description and Analysis.* Reading, Massachusetts, Addison-Wesley Publishing Company.

Howe, Robert W. (1964). *The Relationship of Learning Outcomes to Selected Teacher Factors and Teaching Methods in Tenth Grade Biology Classes in Oregon.* University Microfilms, Ann Arbor, Michigan.

Hoyt, C. (1941). "Test Reliability Established by Analysis of Variance", *Psychometrika, 6,* 153-160.

Kochenderfer, Leonard H. (1966). *A Comparative Study of the Classroom Practice and Teaching Rationale of High School Biology Teachers Using Different Curriculum Materials.* University Microfilms, Ann Arbor, Michigan.

Shaw, Marvin E. and Wright, Jack M. (1967). *Scales for the Measurement of Attitudes.* New York, McGraw-Hill Publishing Company.

AN EXPERIMENTAL STUDY COMPARING A GUIDED DISCOVERY METHOD WITH TWO OTHER METHODS OF TEACHING A PRIMARY SCHOOL SCIENCE UNIT

Rosita Young

Many of the new science curricula developed since 1960 for secondary and primary school pupils stress the methods of inquiry and learning by discovery enthusiastically advocated by Jerome S. Bruner (1960).

Discovery learning has been variously defined. Furthermore, it is commonly assumed that discovery learning allows for errorful learning, that it is guided to some extent, and that it is the outcome of inductive methods of instruction. Reception learning, on the other hand, is generally considered to be associated with the greater degree of guidance inherent in expository deductive methods of instruction. However, Ausubel (1968) proposes two distinct dimensions of discovery-reception learning and inductive-deductive presentation. In this study the term guided discovery is used to refer to an inductive approach (here termed, the discovery-pupil experimentation method) in which errorful learning is gradually reduced.

Despite definitional confusion, it has been widely claimed that discovery learning results in the achievement of many different types of process and product objectives. However, Wittrock (1966) suggests that very few of the claims have been empirically substantiated. Hermann (1969) points out that the results of experiments which have been carried out, mostly of short duration and of the "laboratory" type, are conflicting and often not statistically significant but he concludes that they tend to favour discovery learning.

Some researchers, for example, Thomas and Snider (1969), and Baumgart (1969) have found that inquiry methods favour above average students in some circumstances. Thomas and Snider also found that while a guided-discovery method favoured the acquisition of inquiry skills, factual-conceptual achievement was favoured by a didactic method in (secondary) pupils. Meyer (1970) reports that trial Nuffield Science (inquiry-based) materials were associated with some favourable attitude changes, particularly in girls.

Ramsey and Howe (1969) found confusion and conflicting results in reviewing the few studies carried out with elementary school children, but they suggest that inductive and problem-solving methods seem to bring about more significant achievement. However, Butts and Jones (1966) found that inquiry training was associated with changes in problem-solving behaviour but not with changes in recall of factual information.

Even such proponents of reception learning as Ausubel concede that discovery learning may be relevant to pupils under 12 years but few controlled experimental studies have been carried out with young children in classroom environments.

The benefits which Bruner (1961) attributes to discovery learning may be stated in terms of four general hypotheses:

1 There is a gain in ability to organize information and to apply the organized information in later problem-solving.

2 The child's intrinsic motivation increases and he becomes less reliant on others.

3 Practice in discovery learning results in the learning of the art and technique of discovery.

4 Discovery learning aids retention.

The study reported here was designed to seek more information relating to the preceding general hypotheses by studying a series of questions arising from them.

The Problem

A The primary school children are exposed to a guided discovery approach in the teaching of science; do they demonstrate the cognitive and affective changes consistent with the last mentioned hypotheses? In particular:

1 Is there a gain in the acquisition of understanding of, and of the ability to apply, the concepts, generalizations and principles of science, i.e., is there a gain in the achievement of subject matter?

2 Is there a gain in the ability to use the content-free logical processes of science and the scientific attitudes inherent in the use of inquiry methods?

3 Is there an increased interest in the topics and activities of science?
Is there a more favourable general attitude towards science lessons?
Is there a greater preference for independent pupil-centred activity in science lessons?

4 Is there a gain in the retention of subject matter acquired?

B Are such factors as ability level (measured by intelligence), age, and sex related to these changes?

The Design of the Experiment

In view of the previous discussion it seemed that the questions generated might be appropriately answered by means of an experimental study involving middle primary school children in a classroom setting, with carefully defined teaching methods used over a long period of time, followed by a multivariate analysis of outcomes including the examination of possible interactions. Three instructional techniques were chosen on the grounds of their representativeness among commonly accepted techniques of instruction. They were designated the expository-teacher demonstration method (Method 1), the expository-pupil experimentation method (Method 2), and the discovery-pupil experimentation method (Method 3). Only the latter was considered to be an example of guided discovery.

The experimental design elected was a non-equivalent control-group design classified as quasi-experimental by Campbell and Stanley (1963). The desirable procedure of randomly assigning all subjects to groups was not possible because of the necessity to use three existing class groups which were, however, considered by the school principal, to be parallel for ability. The design was strengthened by the statistical imposition of equivalence by means of analyses of covariance using pre-test and IQ scores as constant covariates.

The Sample

Four coeducational state primary schools from various locations in the Western Metropolitan Area Directorate of the NSW Department of Education were arbitrarily selected. Pupils who had entered fourth grade two weeks previously - of average age 9.1 years - were selected and assigned to learning situations characterized by one of four possible methods (three experimental and one control). The two sets of four sub-groups, derived from four separate schools, were combined to form one set of four treatment groups for the purposes of statistical analysis. Details of the assignation of subjects are shown in Table 1.

TABLE 1

Assignation of Subjects

			Method 1	Method 2	Method 3	Control	
Experimental Sub-Groups	School 1 (Large)	111	37(E)	37(E)	37(E)		
	School 2 (Small)	58	19(R)	19(R)	20(R)		
Control Sub-Groups	School 3 (Large)	118				36(R)	
	School 4 (Small)	54				20(R)	
TREATMENT GROUPS			56	56	57	56	225
	Total	341					Total

Key:
E = existing parallel class group of subjects
R = group of randomly assigned subjects

The Experimental Procedures

The learning sessions consisted of 16 weekly 45-minute classroom lessons which were conducted by a single teacher in successive teaching weeks from the second month of Term 1, 1970. They were preceded and followed by short testing sessions which occurred daily in a two week period before and after the period of the learning sessions. All pre-tests were completed one week before the commencement of the learning sessions. Post-tests of achievement only were administered during the period of the learning sessions.

The instructional unit was devised and tested by the experimenter in fourth grade classes at the end of 1969.

The Independent Variables: Independent variables included ability level, sex and instructional treatment. There was relatively little variation of age among subjects.

Intelligence test scores were obtained from the NSW Department of Education. The median of each treatment group (100 or 101 in all cases) was calculated and used to determine high and low ability groups.

The presence of the three prescribed instructional methods was shown by the use of two measures:

(i) Analysis of audiotapes using Flanders System of Interaction Analysis (Flanders, 1960, p. 20); this technique has been used in science classrooms by Citron and Barnes (1970).

(ii) A Draw-Your-Science Class test devised by the experimenter which was administered as a post-test.

The three instructional methods were used in conjunction with the same subject matter. The sixteen lessons of the instructional unit were divided into four sets of four, each set being based upon a topic selected from one of the four major subject areas of Biology, Chemistry, Geology, and Physics.

Since the topics selected - Germination, Burning, Rocks and Minerals, and Electrical Circuits - concern phenomena familiar to children, and are found in most elementary science curricula including that designed for primary schools in NSW, it was assumed that adequate general background would have been experienced. However, an attempt was made to include relatively unfamiliar subject matter in order to control for possible differences among pupils in this regard.

A number of scientific concepts and generalizations inherent in the topics and suited to the three methods were introduced within the subject matter. These were varied in difficulty within lessons, and lessons were equated within and across topics in this regard. In the explanation of the concepts and generalizations, the logical processes of science and general scientific attitudes were also introduced. The lessons, therefore, involved learning experiences of various types.

Three analagous sets of instructional materials were designed, one appropriate to each instructional method. Each set provided the following items for each of the 16 lessons:

(i) A detailed lesson outline containing a structured sequence of subject matter, including experimental investigations.
(ii) A roughly scripted list of key statements (of concepts and generalizations) and of key questions to be uttered by the teacher.
(iii) A set of overhead projector transparencies containing appropriate diagrams and written statements.
(iv) Sets of individual experimental kits for each lesson, containing simple apparatus (much of which was familiar to children) and experimental materials. In the case of the expository-teacher demonstration method only one experimental kit, containing similar, but larger pieces of apparatus, was prepared for each lesson.

The three experimental methods were prescribed by these three sets of instructional materials and by guidelines as to requisite teacher and pupil behaviours appropriate to each method. (It was assumed that the control group of subjects would experience little, if any, formal systematic instruction in science.)

Two variables were considered to be of basic importance in defining the three instructional methods: firstly, the sequence of presentation of the concept or generalization, principle or rule, and of examples or instances - whether this be an inductive or deductive sequence; secondly, the nature of the guidance received - whether inquiry activities be largely under the control of the teacher or of the individual learner. Other variables considered to be subsumed within the latter include the presence and means of verbalization of the generalization (by pupils or teacher), the means of answering questions, the means of elimination of errors, and the extent of informal guidance made possible by pupil-pupil interaction.

The three instructional methods of expository-teacher demonstration (Method 1), expository-pupil experimentation, (Method 2) and discovery-pupil experimentation (Method 3) were defined and characterized by the properties summarized in Table 2.

In Methods 1 and 2, in each instructional sequence, pupils were encouraged to supply their own answers to questions and to correct errors, but the teacher provided answers and clarified readily. In Method 3 the pupils were questioned increasingly closely in order to lead them to answers to questions and the correction of errors; selection of instances by pupils to illustrate the (unknown) concept, and the verbalization of the concept and relation of instances to it, was similarly carried out initially by pupils as a result of teacher questioning; the statement of the concept was finally repeated by the teacher.

Independent variables such as length of lesson, time of day and the use of recording equipment were carefully controlled. Subjects experienced the same instructional method for the duration of the experiment to avoid any possible confounding due to transfer, which could have arisen had each group of subjects experienced more than one treatment. Transfer from other learning situations in the curriculum could not be controlled. It was assumed that such variables as socioeconomic status, mean group intelligence, reading ability, motivation and attentiveness were equivalent in the sub-groups.

TABLE 2

Resume of Characteristics of Experimental Methods

Property	Method 1	Method 2	Method 3
Presentation of concept	Initially	Initially	Terminally
Presentation of instances	After concept	After concept	Before concept
Verbalisation of concept	By teacher	By teacher	By pupils By teacher
Exploration by experimentation	By teacher	By pupils By teacher	By pupils
Verbal guidance: Questions	Answers given	Answers given	Answers not given
Errors	Shown	Shown	Not shown

The Dependent Variables: The outcomes of the instructional unit were defined in terms of appropriate achievement of subject matter (acquisition and retention), science processes, scientific attitudes, interest in scientific topics, pupil independence in experimental investigations and attitude towards science lessons.

Prior to the experiment pencil-and-paper tests were therefore constructed by the experimenter, refined by item analysis and data relating to their reliability and validity collected. Table 3 illustrates the operationalization of the dependent variables.

The two pairs of post-tests of achievement were administered during the course of the experiment and then only to the subjects experiencing the three experimental treatments. Two of these tests were divided into sections appropriate to individual lessons and administered at the end of these lessons: these were combined as one test of acquisition of subject matter. Two of the tests were administered one week after the completion of the two appropriate topics: they were combined as one test of retention. The achievement tests contained items of varying difficulty and included items testing both understanding and application.

Analysis of the Data

The Experimental Methods: Analysis of verbal interaction was carried out by a trained observer using Flanders' System of Interaction Analysis. Ten minute segments of discourse were arbitrarily selected and ID ratios and other indices thought to be related to verbal guidance were calculated for each method as indicated in Table 4.

TABLE 3

Resume of Dependent Variables and Associated Instrumentation

General Hypothesis	Dependent Variable	Title	Measurement Index: Coefficients of Reliability — Kuder-Rich. 20	Split Half	Test Retest
	Achievement in Subject Matter				
1	Acquisition	(Geol.Ach.Test (Phys.Ach.Test	.685 .783		
4	Retention	(Biol.Ach.Test (Chem.Ach.Test	.763 .806		
3	Processes of Science	Test of Science Processes	.757		
3	Scientific Attitudes	Test of Scientific Attitudes		.684	.566
2	Interest in Scientific Topics	Test of Interest in Science		.721	.959
2	Attitude towards Science Lessons	Test of Attitude towards Science Lessons		.734	
2	Independence	Test of Independence I Test of Independence II		.456 .516	

TABLE 4

(Using Flanders' System of Interaction Analysis)

Indices of Verbal Interaction

	Method 1			Method 2			Method 3		
	ID Ratio	Cat.4 / Cat.5	Cats.8&9 / 100	ID Ratio	Cat.4 / Cat.5	Cats. 8&9 / 100	ID Ratio	Cat.4 / Cat.5	Cats.8&9 / 100
School I Physics	.38	.28	8.25						
School I Geol.				.44	.27	6.70			
School I Biol.							4.45	3.45	17.50.
School I Chem.	.17	.12	5.30	.24	.15	8.85	4.42	4.14	15.65
School II Chem.	.14	.10	8.20	.36	.23	8.40	3.55	2.64	21.50.

From the pictures obtained from the administration of the Draw-Your-Science Class Test, 17 pictures from each treatment group were randomly selected. Two artists, experienced in art work connected with education, independently assigned the pictures to three categories corresponding to the three instructional methods in terms of pupil autonomy in experimentation.

The Criterion Measures: An A x B factorial design was selected for the analysis of data relating to the criterion measures. Analyses of covariance were carried out separately for each of the seven dependent variables using treatment and one learner independent vaiiable in each analysis. Pre-test scores, for each relevant independent variables, and IQ scores were used as covariates in each analysis. When overall significant differences were apparent among treatments but there was no significant interaction with the learner independent variable, individual post-hoc comparisons of treatments were carried out using the technique suggested by Winer for post-hoc comparisons following a one way analysis of variance (Winer, 1962, p. 485).

Results and Discussions

Validation of the Experimental Methods: The indices in Table 4 show a clear difference in the verbal classroom interaction which occurred between Method 3 and Methods 1 and 2. The proportions of pictures assigned by the raters to categories corresponding to the methods experienced by the pupils who drew them were 76 per cent and 70 per cent, an average of 73 per cent. Together these findings indicate that some of the required differences relating to guidance did exist between methods and that pupils perceived them to exist. It was considered that the use of scripts and transparencies with the prescribed inductive and deductive sequences of content controlled its presentation sufficiently closely.

The Criterion Measures: Tables 5-12 show the results of factorial analyses carried out for each of the independent variables. Figures 1-11 show the corresponding profiles where significant results were obtained.

Subject matter information: With regard to the *acquisition* of subject matter, Table 5 and figure 1 indicate that the discovery-pupil experimentation method (Method 3) is the least effective and that the expository-teacher demonstration method (Method 1) is the most effective of the three methods. These effects are even more apparent when the interactions of the methods with high and low ability levels are considered.

With regard to the *retention* of subject matter, Table 6 and figure 2 indicate that the same distinction exists between methods but that there is no significant interaction of methods with ability levels. Differences at the .01 level of significance existed, as indicated below, between Methods 1 and 3 when the adjusted means were compared individually:

(Methods which did not differ significantly from each other are shown on the same line.)

I II
* * * *
* * *
II III

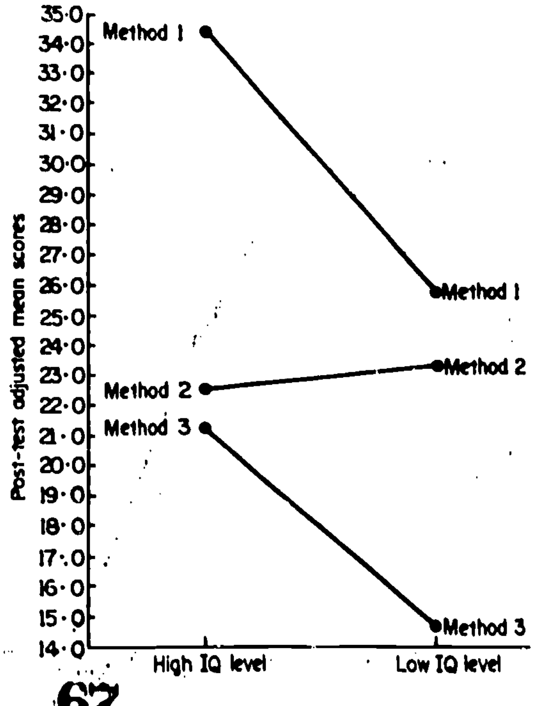
TABLE 5

Test of Achievement (Acquisition)
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.s.	F
1. Between Methods	2	1391.08	63.35**
Between Ability Levels	1	1.02	.05
Interaction Methods x Ability Levels	2	336.52	15.81**
Within	142	21.29	
2. Between Methods	2	1394.91	53.34**
Between Sexes	1	20.45	.78
Interaction Methods x Sexes	2	54.57	2.10
Within	118	26.15	

*Significant at .05 level
**Significant at .01 level

Figure 1



67

TABLE 6

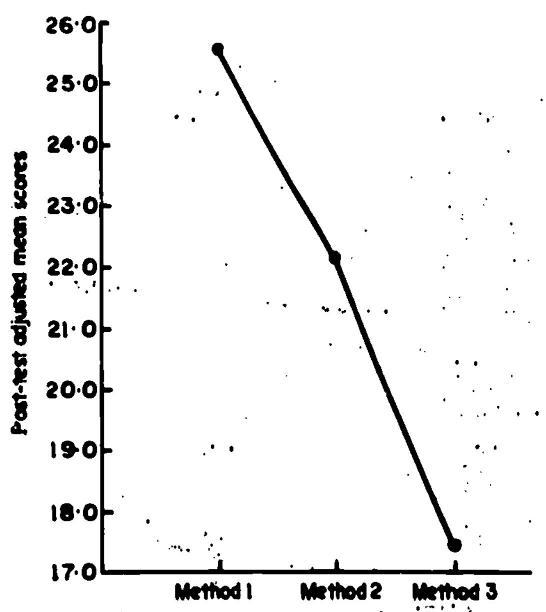
Test of Achievement (Retention)
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.s.	F
1. Between Methods	2	463.67	10.44**
Between Ability Levels	1	130.35	2.44
Interaction Methods x Ability Levels	2	92.13	2.08
Within	142	44.41	
2. Between Methods	2	404.37	8.80**
Between Sexes	1	36.81	.08
Interaction Methods x Sexes	2	40.98	.89
Within	118	45.97	

*Significant at .05 level

**Significant at .01 level

Figure 2



68

Heuristics of discovery: With regard to the *ability* to use science processes, Table 7 and figure 3 indicate that all three methods were superior to the absence of instruction. Method 3 is clearly the most effective and Method 1 the least effective for high ability pupils. It is possible that the test items were not sufficiently discriminatory to distinguish adequately among the three methods for low ability pupils or that the learning tasks were too difficult for them (but as indicated previously other researchers have reported such findings). With regard to the formation of scientific attitudes, Table 8 and figure 4 indicate the same trend in favour of Method 3 for high ability pupils. Method 1 is here clearly the most effective for low ability pupils.

TABLE 7

Test of Science Processes
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.s	F
1. Between Methods	3	355.07	17.92**
Between Ability Levels	1	1732.03	87.39**
Interaction Methods x Ability Level	3	80.36	4.05**
Within	206	19.82	
2. Between Methods	3	237.59	10.12**
Between Sexes	1	12.69	.43
Interaction Methods x Sexes	3	2.92	.01
Within	166	29.41	

*Significant at .05 level
**Significant at .01 level

Figure 3

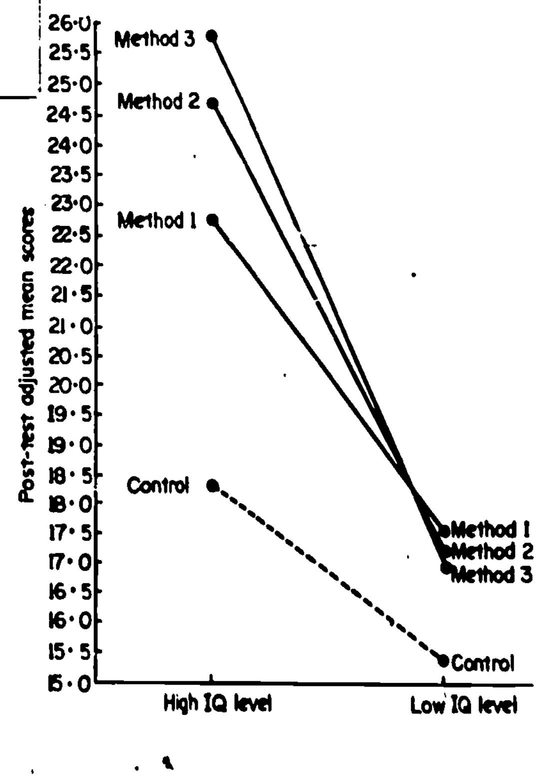


TABLE 8

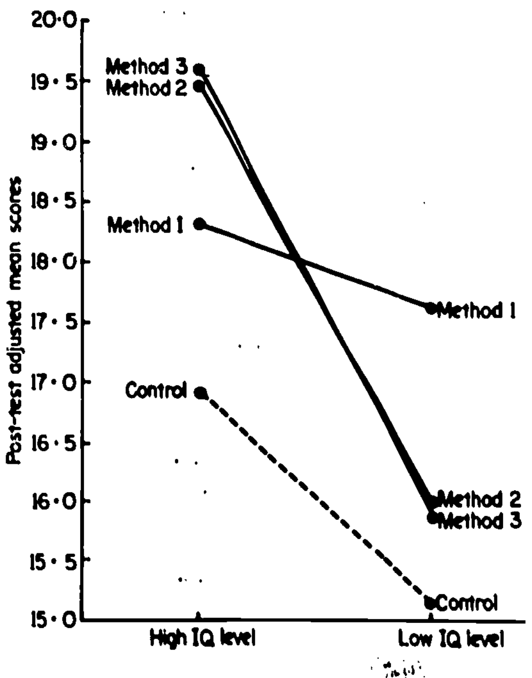
Test of Scientific Attitudes
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.s.	F
1. Between Methods	3	24.96	1.94
Between Ability Levels	1	11.66	.90
Interaction Methods x Ability Levels	3	44.15	3.43*
Within	174	12.89	
2. Between Methods	3	29.38	2.05
Between Sexes	1	14.62	1.02
Interaction Methods x Sexes	3	25.02	1.75
Within	134	14.31	

*Significant at .05 level

**Significant at .01 level

Figure 4



Motivation: With regard to an *interest* in science, Table 9 and figure 5 and 6 appear to indicate that Method 3 is superior to the other methods. However, individual comparisons indicate only that Methods 2 and 3 are significantly superior to the control treatment at the .05 level, as shown below:

(Methods which did not differ significantly from each other are shown on same line.)

III II I
* * * * * *
C I

Method 3 was particularly favourable to girls, who were also particularly affected by the absence of any treatment. This type of finding has previously been interpreted as being in large part due to the novelty of participating in an experiment (Meyer, 1970) but the range of adjusted means for the female sub-groups would tend to suggest that in this situation the opportunity to be involved with, and particularly, to freely carry out scientific investigations, is a major reason for this result.

TABLE 9

Test of Interest in Science
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	m.s.	F
1. Between Methods	3	657.15	11.13 **
Between Ability Levels	1	227.45	3.85
Interaction Methods x Ability Levels	3	1.32	.02
Within	190	59.04	
2. Between Methods	3	587.39	9.92 **
Between Sexes	1	63.06	1.06
Interaction Methods x Sexes	3	207.97	3.51 *
Within	166	59.21	

* Significant at .05 level
** Significant at .01 level

Figure 5

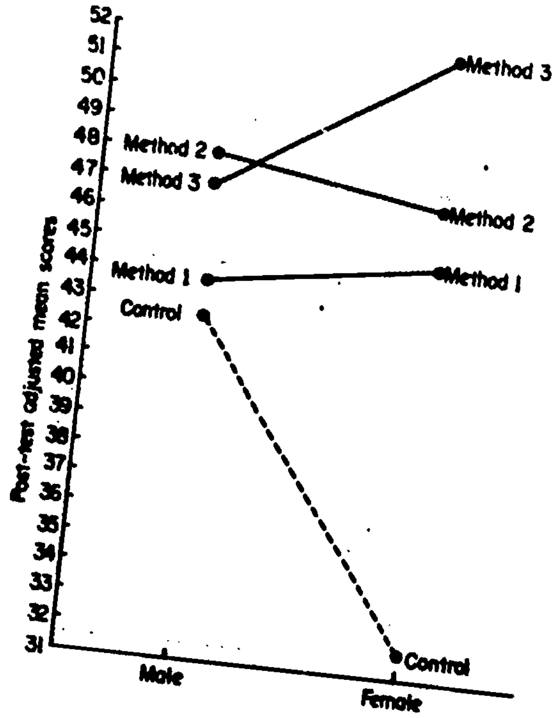
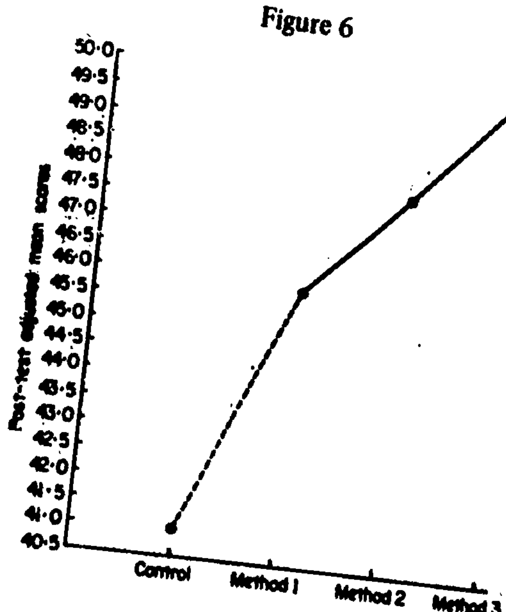


Figure 6



With regard to the test of *attitudes* towards science lessons, Table 10 and figure 7 show that the same overall trend is apparent but here comparisons of means indicate significant differences at the .05 level as shown:

(Methods which did not differ significantly from each other are shown on the same line.)

III II
* * * *
II I
* * * *
I C

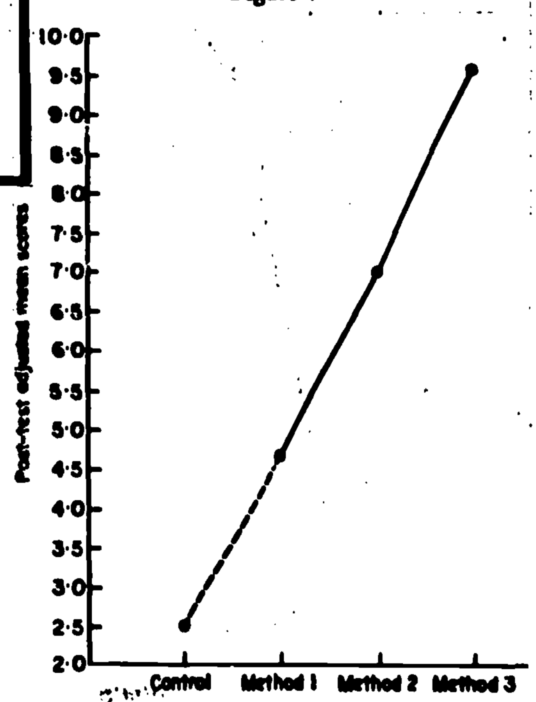
TABLE 10

Test of Attitudes Towards Science Lessons
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	m.s.	F
1. Between Methods	3	639.64	13.37 **
Between Ability Levels	1	35.22	.74
Interaction Methods x Ability Level	3	71.22	1.49
Within	214	47.85	
2. Between Methods	3	444.84	8.93 **
Between Sexes	1	2.29	.05
Interaction Methods x Sexes	3	21.99	.44
Within	182	49.80	

* Significant at .05 level
** Significant at .01 level

Figure 7



73

TABLE 11

Test of Independence I
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.e.	F
1. Between Methods	3	11.80	2.47
Between Ability Levels	1	19.08	3.96 **
Interaction Methods x Ability Levels	3	3.00	.62
Within	214	4.82	
2. Between Methods	3	17.63	3.66 *
Between Sexes	1	3.80	.79
Interaction Methods x Sexes	3	6.56	1.36
Within	182	4.82	

* Significant at .05 level
** Significant at .01 level

Figure 8

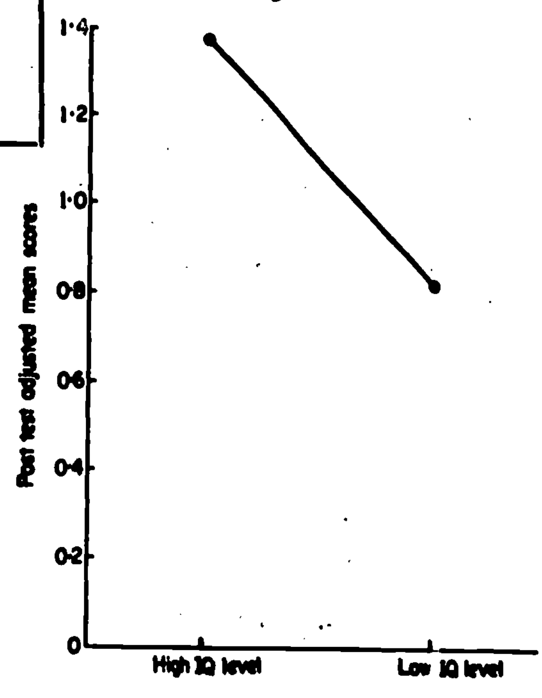
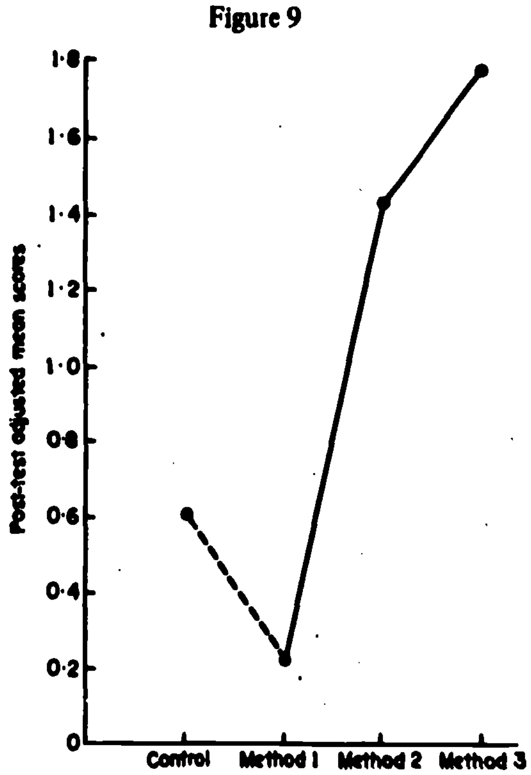


Figure 9



With regard to the tests of *independence* Table 11 and 12 and figures, 8, 9, 10, and 11 indicate that in general the most favourable attitudes are produced by Method 3, particularly with pupils of high ability. It appears from the tables that subjects who experience Method 1 become less independent than those of the control group and that subjects who experience increasingly less guidance become increasingly more independent. The effect is particularly marked for pupils of high ability level (who are also more independent regardless of treatment). This effect is demonstrated to some extent by the individual comparisons:

(Methods which did not differ significantly from each other are shown on the same line.)

Ind. Test 1	Ind. Test 2
III	III
**	**
II	II
**	**
I C	C II
**	**
	I

This finding is perhaps related to that of Flanders (1960) who found that indirect and direct verbal influences were associated with positive and negative attitudes (including dependence proneness) respectively. However, guidance was not varied in a verbal way between Methods 1 and 2 whereas differences in independence are apparent between the two methods in these analyses. It is probable that videotape recordings and the use of an observation system which includes non-verbal caregories suited to science classes, would have discriminated between Methods 1 and 2 with respect to total guidance.

TABLE 12

Test of Independence II
Analysis of Covariance
Comparisons of Post-Test Means

Source of Variation	df	s.s.	F
1. Between Methods	3	22.83	2.96 *
Between Ability Levels	1	69.33	8.99 **
Interaction Methods x Ability Levels	3	23.68	3.07 *
Within	214	7.72	
2. Between Methods	3	33.85	4.14 *
Between Sexes	1	26.11	3.20
Interaction Methods x Sexes	3	13.20	1.62
Within	182	8.16	

* Significant at .05 level
** Significant at .01 level

Figure 10

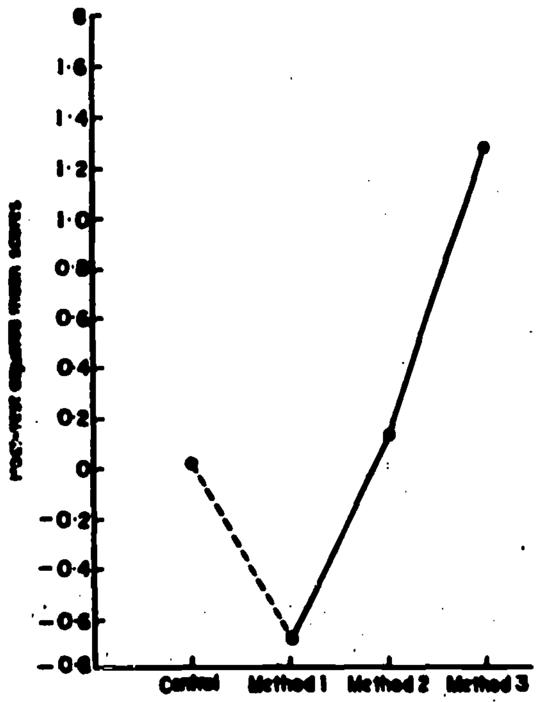
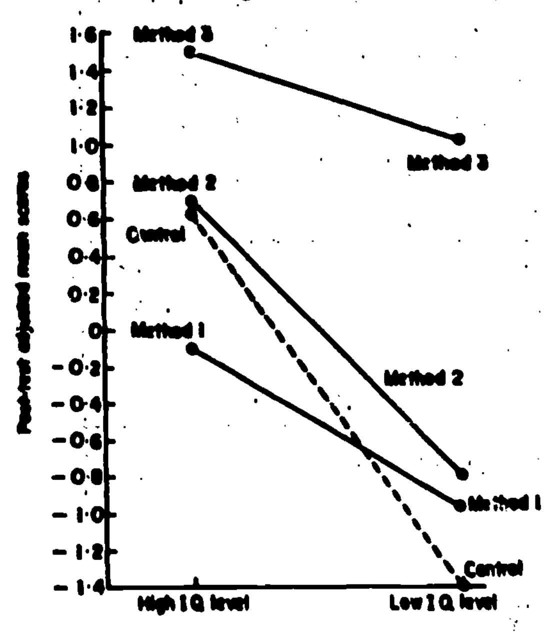


Figure 11



Conclusion

The results obtained tend to support Bruner's claims in relation to the motivation of the learner, since the discovery-pupil experimentation method resulted in the development of an interest in scientific topics, of favourable attitudes towards science lessons, and of a preference for independent activity in science lessons. The claim relating to the heuristics of discovery was supported for high ability pupils only. The claims relating to subject information were not supported.

Weaknesses in the sampling technique have been noted, but it seems reasonable to infer that the findings are applicable to typical fourth grade suburban classes in NSW state schools. A combination of teaching styles should therefore be chosen according to the nature of the objectives and of the learners involved.

REFERENCES

Ausubel, D.P. (1968). *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart and Winston.

Baumgart, N.L. (1969). "The Discovery Approach and Educational Set", *The Australian Science Teachers Journal, 15*, No. 1, 41-50.

Bruner, J.S. (1960). *The Process of Education*, Cambridge: Harvard University Press.

Bruner, J.S. (1961). " The Act of Discovery", *Harvard Educational Review, 31*, 2-32.

Butts, D. and Jones, H. (1966). "Inquiry Training and Problem Solving in Elementary School Children", *Journal of Research in Science Teaching, 4*, 21-27.

Campbell, D.T. and Stanley, J.C. (1963). "Experimental and Quasi-Experimental Design for Research", in *Handbook of Research in Teaching* (Ed. Gage, N.L.), Chicago, Rand McNally & Co.

Citron, J.M. and Barnes, C.W. (1970). "The Search for More Effective Methods of Teaching High School Biology to Slow Learners Through Interaction Analysis; Part 1. The Effect of Varying Teaching Patterns", *Journal of Research in Science Teaching, 7*, 9-19.

Flanders, N.A. (1960). *Teacher Influence, Pupil Attitudes and Pupil Achievement: Studies in Interaction Analysis*, Co-operative Research Monograph No. 12, Office of Education: US Dept. of Health Education and Welfare.

Hermann, G. (1969). "Learning by Discovery: A Critical Review of Studies", *The Journal of Experimental Education, 38*, 58-72.

Meyer, G.R. (1970). "Reactions of Pupils to Nuffield Science Teaching Project Trial Materials in England at the Ordinary Level of the General Level Certificate of Education", *Journal of Research in Science Teaching, 7*, 283-302.

Ramsey, G.A. and How, R.W. (1969) " Analysis of Research Related to Instructional Procedures in Elementary School Science", *Science and Children*, April, 25-36.

Thomas, B. and Snider, B. (1969) "The Effects of Instructional Method upon the Acquisition of Inquiry Skills", *Journal of Research in Science Teaching*, 6, 377-386.

Winer, B.J. (1962). *Statistical Principles in Experimental Design*, United States of America, McGraw Hill Inc.

Wittrock, M.C. (1966). "The Learning by Discovery Hypothesis", in *Learning by Discovery: A Critical Appraisal*, (Eds. Shulman, L.S. and Keisler, C.R.), Chicago: Rand McNally & Co.

SELF INSTRUCTION USING MICROTEACHING AND VIDEOTAPE FEEDBACK DURING AN IN-SERVICE EDUCATION COURSE FOR TEACHERS OF HARVARD PROJECT PHYSICS

J. Rentoul

Introduction: Teaching the "New" Science

Numerous national science projects such as P.S.S.C., the Chemical Bond Approach, B.S.C.S., Nuffield and Harvard Project Physics have devised new materials stressing inquiry as an approach to science instruction. One of the basic aims of Project Physics, for example, is to develop an understanding of scientific processes or methodology along with facts, concepts, and principles. Physics is to be presented as a process rather than as a list of facts to be memorized. Emphasis must be placed on *finding* answers rather than merely on the answers themselves.

Important objectives for students undertaking Project Physics, or any other course based on directed inquiry, are the mastery of skills such as identifying problems, planning experiments, collecting data by observing, measuring and classifying, interpreting and analysing data, drawing inferences, discovering relationships, suggesting hypotheses and making generalizations, predicting, evaluating and questioning information in a critical manner, and identifying fundamental principles and using or applying them in a new situation.

The rationale or philosophy inherent in the "national" courses implies the attainment of certain behavioural objectives by the students which require levels of thinking far in excess of mere memorization and recall of facts. The students should learn to use ideas and factual information rather than simply to remember them.

The attainment of these objectives also necessitates certain teaching behaviours or skills which may elicit the appropriate responses from students. Teachers must stimulate critical thinking and develop an inquiry attitude in their students. The teacher's task is to stimulate investigation and explanation, rather than to give answers and present facts.

It seems essential that children should be encouraged to work at more than one cognitive level and the teacher may assist in the accomplishment of this goal by using careful questioning techniques and by effectively using the students' own ideas.

Many writers including Koran (1969), Wilson (1969), Thiele (1954), Harris (1964), Davis and Tinsley (1967) have stressed the importance of questioning behaviour for the teacher of the new science curricula. In discussing the results of their study, Davis and Tinsley were of the opinion that specific understandings and skills of classroom questioning and the purposes of questions need major attention in the pre-service and in-service education of teachers.

Wilson, in his study of the questioning behaviour of science teachers trained in the inquiry-discovery approach, found that not only were the new science teachers asking more questions of higher-cognitive type, they were asking more questions.

From the above it would seem that one of the basic aims common to the new science curricula is to foster a spirit of directed inquiry in the classroom.

Need for Special Teacher Education

There is considerable evidence from observations and studies to suggest that teachers' behaviour is not of the kind which would seem to be most appropriate for the new science curricula. Studies by Hurd (1966), and Perkes (1967) suggest that the "inquiring style" of teaching which is meant to accompany new materials is little used and poorly understood. Rosenshine (1970) notes, for example, that several investigations, including those by Gallagher (1966, 1968), Parakh (1967, 1968), Harris and Serwer (1966), and Harris *et al.* (1968), report on a great variability of teacher behaviour within specific curriculum. Other investigators, e.g., Evans (1969), Furst and Honigman (1969), Hunter (1969), and Wright (1967) have found few significant differences in the behaviour in traditional classrooms compared to the behaviour in classrooms in which special instructional materials were used. Rosenshine believes that the results obtained may be related to the training programme. In the research cited the teachers were given new materials and new books, but, apparently, they were not given training in the necessary instructional techniques.

New, more efficient and effective methods of teacher training must be devised if new curricula such as Harvard Physics are to be successfully implemented. However, this is only the beginning. It is one thing to train teachers in the required behaviour. It is entirely another problem to ensure that they perform in the classroom in the style required by the new curriculum. We must be certain that not only is there immediate transfer of training to the classroom, but that there is retention over a period of time.

Different Approaches to Teacher Training

Although the need for adequate in-service training for science teachers is widely recognized. there is little agreement as to the form this training should take. Many traditional training programmes for teachers of new courses have stressed the introduction of new content and little attempt has been made to influence the classroom behaviour of teachers by helping them to acquire new teaching skills and techniques.

During the past ten years several new techniques have been studied and developed in an attempt to improve teacher training. Although most of the work has involved the pre-service training of teachers, the following techniques which have evolved can readily be applied to in-service training programmes.

(a) Microteaching and Videotape Feedback

The technique of microteaching and the technical skills approach was developed at Stanford University in 1963. Microteaching focuses on self-evaluation. The teacher may try out a new strategy, or a particular teaching skill, on a limited topic for a short time (5-10 minutes), with a small group of (3-5) students. He can evaluate his strategy as he views the lesson on videotape. Microteaching usually incorporates a supervisor, or consultant, who discusses the teacher's performance with him as they view the playback of the lesson.

Research evidence presented by Borg *et al.* (1968, a & b) indicates that there is a significant transfer of skills, learned in the microteaching situation, to the regular classroom, and that there is a retention of these skills over a period of serveral months.

A comprehensive discussion of microteaching is given by Gage (1968) and Allen and Ryan (1969).

(b) Perceptual Modelling

Another training approach which is currently receiving considerable attention is the use of model videotapes of experienced teachers performing a short lesson to illustrate the use of a particular skill.

(c) Interaction Analysis

This is a technique for observing and analyzing teacher and pupil verbal and non-verbal behaviour in the classroom.

A number of studies have shown that teachers given training in interaction analysis have been able to modify some aspects of their classroom verbal behaviour. Examples of research of this kind used within teacher training courses are provided by Kirk (1967), and Hough and Ober (1967).

Related Research

There is sufficient evidence from both empirical studies and anecdotal reports from those who have participated in teacher training programmes to indicate that microteaching, skill training, and video recording techniques have definite potential in the field of teacher education.

The following table (Table 1) summarizes some of the methods of presentation, feedback and the ways in which supervisors have been used in recent studies. Combinations of these procedures have been employed in studies by Orme (1966), McDonald and Allen (1967), Berliner (1969), and Claus (1969).

Research findings from experimental designs incorporating various combinations of presentation and feedback (as in Table 1) seem to suggest that some skills can be acquired and teaching behaviours modified by using carefully matched model tapes and by allowing the trainee to view his own classroom performance. This last mentioned technique has been found to be more powerful if reinforcement and a contingent focus are provided by a trained supervisor, especially during the presentation of a model. Evidence exists that for some skills symbolic modelling (written instructions or transcript of model lesson) could be used rather than expensive perceptual models.

TABLE I

Summary of Techniques Employed in Microteaching Situations

Type of Presentation	Type of Feedback	Purpose of Supervisor.	Period of Supervision.
1. Perceptual Model (showing what to do) Videorecording. 2. Symbolic Model (Telling what to do) Written instructions OR Transcript of a model lesson.	1. View own performance alone, without cue discrimination training or reinforcement. 2. View own performance alone, with cueing and/or reinforcement provided by (a) written instructions (b) commentary on parallel sound track. 3. View own performance with Supervisor, providing cueing and/or reinforcement. 4. View model alone (prompting feedback) with or without cueing provided by written instructions or sound track commentary. 5. View model with supervisor providing cueing. 6. Feedback from symbolic model. (confirmation feedback) with or without supervision. 7. View self and view model. (a) Alone or with Supervisor. (b) With or without cueing and/or reinforcement.	1. Cue discrimination training 2. Reinforcement. 3. Cueing and reinforcement.	1. During presentation only 2. During feedback only. 3. During presentation and feedback.

Aims of the Current Investigation

Very few studies have been conducted to ascertain the effectiveness of microteaching and video recording techniques for the in-service training of teachers, in particular, for the training of teachers who are involved with the new science courses.

This investigation attempts to ascertain the feasibility of using self instruction techniques involving microteaching and videotape feedback during in-service education to foster skills thought to be important components in Harvard Project Physics. Specifically, the aims of the study were to:

(i) investigate to what extent teacher behaviour can be modified using techniques which will be described subsequently;

(ii) investigate the relative effectiveness of several methods of self instruction for the acquisition of two teaching skills;

(iii) evaluate teacher reaction to the use of this type of training during an in-service course.

The two teaching skills that were to be acquired by teachers in this study were (i) probing behaviour, and (ii) fluency in questioning. Probing behaviour attempts to lead students to a more comprehensive or thoughtful answer instead of the teacher immediately supplying the correct answer or offering his own ideas and opinions.

Fluency in questioning refers to the use of a wide range of questions which requires the students to operate at different cognitive levels. These particular skills were chosen for reasons presented elsewhere (Rosenshine, 1969).

The term "self-instruction" used in the stated aims of the study refers to the receipt of the presentation and the feedback without the assistance of a supervisor. The teacher receives the presentation and/or playback and evaluates his own performance.

Rather than use a supervisor to provide cues and reinforcement, the self-view condition in this study was combined with a contingent focus in the form of an interaction analysis coding system which forced the trainee to concentrate on specific verbal behaviours while viewing a video recording of his own performance, or the performance of a model.

Statement of Hypotheses

(a) It was hypothesized that significantly greater changes in behaviour would be produced by techniques which provided feedback, than by those which did not include any form of feedback.

(b) It was hypothesized that significantly greater changes in behaviour would be produced by the technique which provided feedback in the form of a perceptual model with a contingent focus rather than feedback consisting of self-view with a contingent focus.

(c) It was hypothesized that there would be significantly greater changes in behaviour displayed by the group which received only symbolic model presentation, than by the control group.

Description of Study

The 45 teachers who were participating in the in-service course for Harvard Project Physics at Macquarie University during December 1970 were randomly assigned to a control group and three treatment groups.

First Day

Teachers in all groups taught a 5-minute lesson on a specified topic to a micro-class of four students (two boys and two girls). Each teacher had two days prior to the first microteaching session in which to prepare the lesson for which he was given certain objectives. The lessons were taught in rooms equipped with videotape equipment operated by a technician. The teacher, the technician and the four students were the only ones present in the room during the 5-minute session. The students in the micro-classes came from a nearby High School and the classes were randomly assigned to teachers and treatment groups.

Table 2 summarizes the treatments for each group.

TABLE 2

Summary of Steps in Treatment by Experimental Group

Steps in Treatment	Experimental Groups (1)	(11)	(111)	(1V)	Minutes in Treatment
Teach (Pre Test)	x	x	x	x	5
Set Induction	x	x	x	x	15
Instructions		x	x	x	
View Self and Code			x		5
View Model and Code				x	5
Re-Teach (Post Test)	x	x	x	x	5
View Pre-Teach after Re-Teach (Not part of experimental treatment)	x	x			5

Following the "pre-test" microteaching, the various steps in treatment were:

(1) Set induction, which included a brief discussion of the main objectives of Harvard Project Physics and the directed inquiry approach.

(2) Instructions for Groups (II), (III), and (IV) were discussed, and emphasized the importance of two teaching skills, namely (a) fluency in questioning, and (b) probing behaviour. The instructions suggested that the teacher should attempt to incorporate these behaviours in his second lesson.

(3) Group (III) viewed and coded their own lesson using a simple coding sheet with categories called fact question, thought question, probing behaviour and informing behaviour.

(4) Group (IV) viewed and coded a model videotape of a teacher emphasizing the above teaching skills.

(5) Each teacher in each group had a choice of three similar lessons for his 5-minute re-teach, partly for the purpose of avoiding comparisons between teachers when they viewed the lessons in groups on the second day, and partly because it provided more variety for the microteaching classes. The re-teach was performed before a different class of students.

(6) Groups (I) and (II) had the opportunity of viewing their own pre-teach after the completion of their second lesson so that all subjects were given the opportunity of viewing their first lesson.

Second Day

The teachers attended the same rooms in groups of three for viewing and discussion of their re-teach from the previous day. This session lasted 1½ hours. It was suggested that they might like to discuss the use of techniques such as questioning and using students' responses.

After discussion each teacher had the opportunity to re-teach his previous lesson which he could then view and discuss again with the group.

After the second microteaching session each participant was asked to complete a questionnaire. The purpose of the questionnaire was to determine teacher reaction to the use of the type of training provided.

Procedure for Testing Hypotheses

In this study seven ratios and four categories were used to measure the acquisition of the teaching skills and hence to test the hypotheses.

The ratios were:

(1) $\frac{\text{Probing}}{\text{Non-Probing}}$ $\frac{(P)}{(NP)}$

(2) $\frac{\text{Probing}}{\text{Total Pupil Talk}}$ $\frac{(P)}{(TPT)}$

(3) $\frac{\text{Probing}}{\text{Total Teacher Talk}}$ $\frac{(P)}{(TTT)}$

(4) $\frac{\text{Probing Questions}}{\text{Total Questions}}$ $\frac{(PQ)}{(To.Q)}$

(5) $\frac{\text{Thought Questions}}{\text{Total Questions}}$ $\frac{(TQ)}{(To.Q)}$

(6) $\frac{\text{Fact Questions}}{\text{Total Questions}}$ $\frac{(FQ)}{(To.Q)}$

(7) $\frac{\text{Fact Questions}}{\text{Thought Questions}}$ $\frac{(FQ)}{(TQ)}$

The categories were:

(8) Probing (P)

(9) Probing Questions (PQ)

(10) Though Questions (TQ)

(11) Fact Questions (FQ)

(12) Total Questions (To.Q)

Ratios (6) and (7) and Category (11) were later omitted from the study when the data were analyzed because of the very small number of fact questions asked by any subject. Ratios (1), (2), (3), (4), and Categories (8), (9) were concerned with Probing Behaviour and the other ratios and categories were related to Fluency in Questioning.

Hypotheses in Operational Terms

I There will be no differences between the combined feedback groups (III and IV) and the combined non-feedback groups (I and II) for the following adjusted ratios and categories:

1 $\frac{P}{NP}$ 2 $\frac{P}{TPT}$ 3 $\frac{P}{TTT}$ 4 $\frac{TQ}{To.Q}$

5 $\frac{PQ}{To.Q}$ 6 TQ 7 P 8 To.Q 9 PQ

II There will be no differences between Group III (Self-view and code), and Group IV (Model-view and code), for the following adjusted ratios and categories:

1 $\frac{P}{NP}$ 2 $\frac{P}{TPT}$ 3 $\frac{P}{TTT}$ 4 $\frac{TQ}{To.Q}$

5 $\frac{PQ}{To.Q}$ 6 TQ 7 P 8 To.Q 9 PQ

III There will be no differences between Group I (Control), and Group II (Symbolic model presentation), for the following adjusted ratios and categories:

1 $\frac{P}{NP}$ 2 $\frac{P}{TPT}$ 3 $\frac{P}{TTT}$ 4 $\frac{TQ}{To.Q}$

5 $\frac{PQ}{To.Q}$ 6 TQ 7 P 8 To.Q 9 PQ

In each case the post-test measure was adjusted using the pre-test measure as the covariate.

Recording and Analysis of Data

Each pre-teach and re-teach lesson from the first day was analysed using a category system similar to that used by the teachers in Groups III and IV except that Probing was subdivided into Question and Clue. Because of the unequal length of lessons the tallies in each category were converted into a rate (number per minute). Due to drop out and technical difficulties only six tapes from each group were analyzed. The reliability of the coding was checked by an independent rater and a Scott coefficient of 0.90 was obtained for the comparison of the coding of 236 teacher questions and statements.

Because Groups I and II were to receive very different treatments from Groups III and IV which were to receive feedback in the form of viewing and coding, and because Groups III and IV differed in the type of feedback (viewing self and viewing model), it was decided to use the technique of planned or *a priori* comparisons to compare the adjusted means of:

(i) Groups I and II;
(ii) Groups III and IV;
(iii) Combined Groups I and II and combined Groups III and IV.

This technique has been used by Winer (1962) and by Hays (1963). The results obtained for the planned comparisons are shown in Table 3 and only those involving significant F ratios are included.

TABLE 3

Summary of Results Obtained for Planned Comparisons[a]

	Source of Variation		S.S.	df	M.S.	F
P / T.T.T	Comparison	1	0.0090	1	0.0090	0.6122
		2	0.0135	1	0.0135	0.9184
		3	0.1476	1	0.1476	7.3069 *
	Adjusted Error	1	0.0147	19		
		2	0.0147			
		3	0.0202			
P	Comparison	1	0.0926	1	0.0926	0.1340
		2	0.2293	1	0.2293	0.3193
		3	7.5978	1	7.5978	7.1956 *
	Adjusted Error	1	0.6908	19		
		2	0.7182			
		3	1.0559			
To.Q	Comparison	1	0.0063	1	0.0063	0.0075
		2	1.0806	1	1.0806	1.1156
		3	10.1850	1	10.1850	10.6449 **
	Adjusted Error	1	0.8404	19		
		2	0.9686			
		3	0.9568			
P.Q.	Comparison	1	0.0132	1	0.0132	0.0273
		2	0.5415	1	0.5415	1.1195
		3	8.5224	1	8.5224	12.5127 **
	Adjusted Error	1	0.4829	19		
		2	0.4837			
		3	0.6811			

* $p < .05$
** $p < .01$

[a] Comparison 1 Group I with Group II
2 Group III with Group IV
3 Combined Groups I and II with Combined Groups III and IV

A further analysis suggested by the results of the planned comparisons was carried out on the data but the results are not reported in this paper. A 2 x 2 factorial design was used in which Groups I and II were combined and Groups III and IV were combined, and both combined groups were dichotomized into high and low levels on the basis of the pre-teach scores for each of the above ratios and categories.

Discussion of Results of Planned Comparisons

There were four significant results for comparison 3 in which the combined non-feedback groups (I and II) were compared with the combined feedback groups (III and IV).

Significant F ratios were obtained for the ratio $\frac{\text{Probing}}{\text{Total Teacher Talk}}$ and for the three categories: Probing, Total Questions and Probing Questions, the last two being significant at the 1 per cent level.

The results indicate that a teaching skill such as Probing can be acquired and teacher behaviour can be modified using a technique which employs a symbolic presentation (written instructions), together with feedback and a contingent focus (view self or model with coding). The provision of feedback proved to be a more important factor than the type of feedback. Written instructions alone were not significantly effective and although there was no significant difference, results tended to favour the technique of prompting feedback provided by model viewing and coding over that of self-viewing and coding.

Since teacher verbal behaviour was modified and a skill acquired in such a short session of microteaching without the presence of a supervisor, it seems that a self instruction technique such as that employed in Groups III and IV may be useful for the in-service education of teachers.

Before the data were obtained it was intended that Fluency in Questioning should be measured by using the ratio $\frac{\text{Fact Questions}}{\text{Thought Questions}}$.

The written instructions suggested that subjects should try to use "a good balance of fact and thought questions" in initiating discussion, and it was hoped that the teachers who were more fluent in their questioning would obtain values nearer unity for the above ratio than those who were less fluent. However, the frequency of Fact Questions was so low for all participants that it was impossible to statistically analyze the data obtained or to compare the ratios of $\frac{\text{Fact Questions}}{\text{Thought Questions}}$ for any of the experimental groups.

It may have been too much to expect that both skills (fluency and probing) could be learned in the time available. In attempting to ask probing questions which were almost invariably questions requiring thought from the students, the teacher may have found it difficult to also employ a number of Fact Questions. Another possibility is that the type of lesson that the participants were asked to teach made the asking of Fact Questions more difficult.

However, there were significant differences obtained for probing questions and for total questions asked. This result tends to agree with the statement made by Claus(1969) that "it may be possible to cause a change in frequency of a behaviour in a short time but a longer time may be needed to change the type of behaviour using these techniques".

REFERENCES

Allen, D.W. and Ryan, K.A. (1969). *Microteaching*, Reading, Mass., Addison-Wesley.

Berliner, D.C. (1969). *Microteaching and the Technical Skills Approach to Teacher Training.* Technical Report No. 8, Standford Centre for Research and Development in Teaching.

Borg, W.R. *et al.* (1968a). *The Minicourse: Rationale and Uses in the Inservice Education of Teachers.* Paper presented at the Annual Meeting of the American EDucational Research Association, Chicago.

Borg, W.R. *et al.* (1968b). *The Minicourse: Changing Teacher Behaviour Through Microteaching,* Unpublished Manuscript, Far West Laboratory for Educational Research and Development.

Clause, K.E. (1969). *Effects of Modeling and Feedback Variables on Questioning Skills.* Technical Report No. 6, Stanford Centre for Research and Development in Teaching.

Davis, O.L. and Tinsley, D.C. (1967). "Cognitive Objectives Revealed by Classroom Questions Asked by Social Studies Student Teachers". In Hyman, Ronald T. (ed.), (1968), *Teaching: Vantage Points for Study*, Philadelphia, J.B. Lippincott.

Evans, Thomas P.A. (1969). "A Category System for Teacher Behaviours", *American Biology Teacher*, 31, 221-26, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res., 40*, No. 2, 279-99.

Furst, Norma and Honigman, Fred. K. (1969). *A Study of Pupil-Teacher Interaction in Experimental Communications Courses Compared with Traditional Junior High School Classes.* Paper presented to the American Educational Research Association, Philadelphia, Temple University, College of Education, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res., 40*, No. 2, 279-99.

Gallagher, James J. (1966). *Teacher Variation in Concept Presentation in BSCS Curriculum Program.* Urbana, Institute for Research on Conceptional Children, Univ. of Ill., cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res., 40*, No. 2, 279-99.

Gallagher, James J. (1968). *Analyses of Teacher Classroom Strategies Associated with Student Cognitive and Affective Performance*, Cooperative Research Project No. 3325, (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res., 40*, No. 2, 279-99.

Harris, Albert J. and Serwer, B. (1966). *Comparison of Reading Approaches in First Grade Teaching with Disadvantaged Children.* Cooperative Research Project No. 2677, New York: City Univ. of New York, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res., 40*, No. 2, 279-99.

Harris, Albert J. *et al.* (1968). *A Continuation of the CRAFT Project: Comparing Reading Approaches with Disadvantaged Negro Children in Primary Grades*, Cooperative Research Project No. 5-0570-2-12-1. New York: Division of Teacher Education of the City Univ. of New York, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res. 40*, No. 2, 279-99.

Hays, William Lee (1963). *Statistics for Psychologists*, pp. 462-489, New York: Holt, Rinehart and Winston.

Hough, John and Ober, Richard (1967). "The Effects of Training in Interaction Analyses on the Verbal Teaching Behaviour of Pre-Service Teachers". In *Interaction Analysis: Theory Research and Application*, Amidon and Hough (eds.) Reading, Mass., Addison-Wesley.

Hunter, Elizabeth (1969). "Talking in First Grade Classrooms", *Urban Review*, 4, 39-43, cited hy Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res.*, *40*. No. 2, 279-99.

Hurd, Paul de H. (1966). "Teacher Preparation in Biology: Programs, Plans, Progress", *Newsletter*, *28*, B.S.C.S.

Kirk, Jeffrey (1967). "Elementary School Student Teachers and Interaction Analyses", In *Interaction Analysis: Theory Research and Application*, Amidon and Hough (eds.), Reading, Mass, Addison-Wesley.

Koran, J.J. (1969). "Two Paradigms for the Training of Science Teachers using Videotape Technology and Simulated Conditions", *J. of Res. in Science Teaching*, *6*, No. 1, 22-28.

McDonald, F.J. and Allen, D.W. (1967). *Training Effects of Feedback and Modeling Procedures on Teaching Performance.* Technical Report No. 3, Stanford Centre for Research and Development in Teaching.

Orme, M.E.J. (1966). *The Effects of Modeling and Feedback Variables on the Acquisition of a Complex Teaching Strategy.* Doctoral dissertation, Stanford University, Ann Arbor, Michigan: University Microfilms, No. 67-4417.

Parakh, Jal S. (1967). "A Study of Teacher-Pupil Interaction in High School Biology Classes: Part II. Description and Analysis", *J. of Res. in Science Teaching*, *5*, 183-92, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res.*, *40*, No. 2, 279-99.

Parakh, Jal S. (1968). "A Study of Teacher-Pupil Interaction in BSCS Yellow Version Biology Classes", *American Biology Teacher*, *30*, 841-48, cited by Rosenshine, Barak (1970), "Evaluation of Classroom Instruction", *Rev. of Educ. Res.*, *40*, No. 2, 279-99.

Perkes, Victor A. (1967). "Junior High School Teacher Preparation, Teacher Behaviours, and Student Achievement", Unpuhlished Doctoral Dissertation, Stanford University.

Rosenshine, Barak (1969). "Teaching Behaviours Related to Pupil Achievement", *Classroom Interaction Newsletter*, *5*, No. 1, 4-17.

Rosenshine, Barak (1970). "Evaluation of Classroom Instruction", *Rev. of Educ. Res.*, *40*, No. 2, 279-99.

Thiele, C.L. (1954). "Fostering Discovery with Children", *Arithmetic Teacher*, *6*, February.

Wilson, John H. (1969). "The 'New' Science Teachers are asking More and Better Questions", *J. of Res. in Science Teaching*, *6*, No. 1, 49-53.

Winer, B.J. (1962). *Statistical Principles in Experimental Design*, p. 208, New York: McGraw-Hill.

THE EFFECTS OF PUPIL INVOLVEMENT IN CLASSROOM INTERACTION OF SCIENCE ACHIEVEMENT

C.N. Power

There is general agreement among science educators that learning is an active process. Most reject the conception of teaching in which the teacher is the remote provider and the pupil the dyspetic receiver. Most accept the principle that active student participation in classroom activities is a necessary condition for learning. Activity, participation, and involvement cannot always be inferred from the superficial appearance of the instructional setting. Some students listening to a meaningful verbal exposition by the teacher may be very active participants; others, despite their exposure to "discovery experience" in laboratory settings, may be passive.

General agreement on the active nature of the learning process should not be construed also to imply agreement upon the desirability of overt response-making, as in a linear programmed text, or of the desirability of direct teacher-pupil interaction. Siegel (1967), for instance, suggests that, whereas frequent direct interaction may facilitate the performance of certain kinds of students, it may not prove optimal for others. Siegel and Siegel (1967) found that televized lectures inhibited conceptual acquisition by students who were initially poorly motivated, unsophisticated in the subject matter area and predisposed to learn factual material rather than to apply, or to systhesize. Personal contact did not influence the criterion performance of those initally sophisticated, or of those conceptually set.

Parallel results have been obtained in studies of the differential effects of programmed instruction on pupils. For instance, the performance gap between person-oriented pupils and task-oriented pupils tends to be greater under instruction from programmed materials than from warm teachers. Requiring pupils to make overt responses in the typical linear programme represents an advantage only for the less able, unsophisticated student.

This decade has seen the advent of more individualized science curricula, where personal contact between teacher and pupil is enhanced. Perhaps, as Gage and Unruh (1967) claim, the time is ripe to re-examine assumptions about what goes on in the conventional classroom. Gage and Unruh go on to ask, "why is it that programmed instruction has not yielded more dramatic results in terms of improving learning?" Perhaps active participation is occurring when a pupil is listening to classroom discussion. Certainly there is a need for careful analyses of who gets what out of teacher-pupil interaction.

An important question is whether greater direct involvement necessarily leads to greater achievement gains. One way of seeking an answer to this question is to explore the differences between groups of students for whom varying levels of participation in classroom interaction does, or does not, lead to higher achievement.

The search for answers to the last mentioned question formed part of a larger study of the effects of communication patterns and teacher feedback in grade 8 science classrooms. In this study video-tapes were taken of five lessons in each of four classrooms, and the teacher and pupil behaviour classified. By-products of the classification were measures of the level of direct participation by each pupil. In addition, measures of the following were obtained -

(a) pupil cognitive abilities
(b) pupil personality
(c) expectation of success in science
(d) achievement motivation
(e) sociometric status
(f) attitude towards science
(g) pre- and post-test measures of science achievement.

Using the measures, pupils were divided into the following six groups depending on whether they were frequently, moderately frequently, or rarely, directly involved with the teacher and whether their residual achievement scores (calculated by regression techniques) were high or low:

Group 1 Students frequently involved who made high gains in achievement
Group 2 Students frequently involved who made low gains in achievement
Group 3 Students occasionally involved who made high gains in achievement
Group 4 Students occasionally involved who made low gains in achievement
Group 5 Students rarely involved who made high gains in achievement
Group 6 Students rarely involved who made low gains in achievement.

Details of the grouping procedure appear elsewhere (Power, 1971).

A multiple discriminant analysis was carried out to determine whether various groups of students within the classroom could be distinguished by a set of variables operating together. Details of this technique can be found in Veldman (1967), and Cooley and Lohnes (1971). The results of the analysis are listed in Appendix I. The lambda coefficient for the six groups was 0.116 which was significant beyond the 0.001 level. This means that the chances of producing group differences as large or larger by drawing six random samples from a 25 dimension multivariate swarm (25 variables were used in the discriminant analysis) were less than 1 in 10,000.

The generalized multivariate null hypothesis that the six groups are similar in personality, ability, expectations of success, sociometric status and attitude towards science was thus regarded as untenable. It was possible, therefore, to proceed to examine the group differences obtained from the group mean vectors, the discriminant weights and correlations, and the group centroids.

Two of the five discriminant functions were found to be highly significant and a third moderately so ($p = 0.08$). These functions represent orthogonal (independent) dimensions which distinguish among the groups. The three functions account for 43.55, 24.33, and 15.36 per cent of the variance respectively.

Function I. Achievement Syndrome

The discriminant weights for this function indicate that by far the largest contribution to group separation along the dimension is contributed by expectation of success, with smaller contributions by ability, achievement motivation, positive expansiveness, and attitude towards science. High scores on these variables result in high scores on Function I. The relatively high means of groups 1 and 3 and the low mean of group 6 on each of these variables largely explains the location of these groups along this dimension.

The function was interpreted as an achievement syndrome characteristic of very successful and involved students. Those high on this dimension need, expect and achieve success. Because of their potential and need dispositions, these pupils probably exert considerable "psychological force" on the teacher which leads to initial involvement. Continuing success during interaction helps to establish them as desirable models as well as to meet their own needs, those of the group and of the teacher. The effects of successful involvement for such students include high expectation of success, high status, and positive attitudes. Higher achievement undoubtedly springs from the greater opportunities for reinforcement, confirmation and clarification which involvement brings.

At the other extreme, the low gain, low involved group are less able, have lower needs for achievement and are less fluent. Consequently, they do not expect success and do not look for need satisfaction in the academic area. Their status is low, attitudes poor, and they tend to be ignored and/or rejected by the teacher and other groups. This non-involvement restricts chances for learning, and increases frustration and boredom.

Function II. Social Orientation

This dimension distinguished best among groups making high gains but differing in the extent of involvement (i.e., groups 1, 3, and 5). Roughly equal positive loadings are contributed by such variables as expectation of success, fluency, sociability (A), surgency (F), sensitivity (I), insecurity (O), and self-control (Q3), with negative contributions by conscientiousness (G⁻), and group dependence (Q2⁻). The high negative loading on Inter D (scholastic aptitude) and the positive loading on brightness (B) are difficult to explain, but the net effect is to lower group 5 relative to groups 1 and 3.

The function scores correlate moderately with variables Q2 (-.36), I (.33), D (-.32), and A (.31), i.e., with group dependences, sensitivity or attention seeking, stodginess and sociability. Thus, the major differences between groups 1, 3, and 5 are related to scores on the dimension which is linked largely with personality characteristics relating to group dependence. The members of group 1 achieve largely within the social context while group 5 pupils achieve largely by vicarious and independent means. Group 3 pupils occupy an intermediate position. Groups 2, 4, and 6 (low gain groups which differ in involvement) tend to be high on this dimension. Groups 2 and 4 are particularly "group dependent". It is possible that members of group 2 and 4 are dependent on the teacher and group 6 on their peers within the group.

Function III. Social Maturity

Groups 2, 4, and 5 have larger negative values on this dimension than the other groups. Variables contributing most to scores on this function are sociometric scores while ability and two personality factors (C and Q2) help differentiate three of the groups, possibly because of their effects on the perceptions of group members. The interrelation among these variables is intriguing. Some comments follow.

The high means of groups 1 and 3 on positive expansiveness and model status tend to ensure them a higher centroid on this dimension, although the value is reduced somewhat by high means on ability and choice status. Group 2 has relatively low means on model status and ego strength (C). Groups 4 and 5 have relatively low means on positive expansiveness, model status and rejection status while group 4 has a low mean on C, a factor which militates against effective group integration. Group 5 has a high mean on Q3 (Group independence).

One possible but highly tentative interpretation of the data is that groups 1, 3, and 6 fulfill their role expectations. Members of groups 1 and 3 are bright, successful and play the expected role of learning models. Group 6 members are academically less able and less successful, and are not expected by their peers to be involved in the direct communication system - an expectation they live up to!

Group 5 is composed of fairly able, but group-independent (Q2) students who prefer to work alone and yet who achieve success. That such students receive more choices and fewer rejections than their involvement would seem to indicate, suggests that other students perceive that they could contribute more to the functioning of the group than they do. On the other hand, group 2 members seem to be perceived as being involved more often than their ability, success or status warrant. Members of this group also tend to be more emotionally unstable (C) and aesthetically sensitive (I).

Group 4 members make fewer choices and more reflections. Like group 2 members they tend to be more emotionally unstable than members of other groups. It would seem that members of these groups exhibit less social maturity in the student role and more often show signs of emotional instability.

Analysis of Variance

The picture may be embossed a little when the results of several three-way analyses of variance are considered. In these ANOVA, scores for residual knowledge, comprehension, application and total science achievement scores were used as criteria. In each case the first factor was the class; the second, high, middle or low frequency, the third, was high or low anxiety, extroversion or ability. The results are summarized in Appendix 11.

Fairly consistent main effects were found on each criterion for class, ability and extroversion, with extroversion inhibiting performance. Main effects appeared for frequency when application and total gain scores were the criteria. The results were interpreted as indicating that *frequent involvement* helps pupils to apply their knowledge better. Interaction effects were found between anxiety and frequency, ability and frequency, class and ability, and class and frequency. Both bright and dull students were found to benefit when involved but average students performed better when they are rarely involved. Anxious students made greater gains when moderately, or rarely involved, while non-anxious students performed better when frequently, or moderately involved, but in the fourth, moderately involved students performed far worse than those often or rarely involved.

Discussion

In essence, the analyses provide a picture of the conditions under which overt participation does, or does not, facilitate science achievement. For the socially oriented, dependent, less able and poorly motivated pupil, being ignored guarantees failure, frustration, withdrawal and eventual alientation from school. Moderate and frequent involvement permits compliance, and a minimal level of performance - but high involvement, for the dependent prone, carries with it the price of rejection by peers, presumably because the high participation hinders group performance.

The task-oriented, bright, motivated students fall into two groups: the non-anxious who enjoy the limelight that successful involvement and leadership bring; and the introverted, anxious, serious students who prefer to work alone. Both groups achieve if congruity exists between their predispositions and the conditions of learning.

If the findings from this study are a guide, there is no doubt that there is an urgent need to create materials and environments which provide for greater pupil participation and involvement in learning tasks. Furthermore, more research is required on the issue of creating environments which facilitate learning in different pupils. There is, too, a need to discover the conditions under which self-esteem, expectation of success, strivings for excellence, independence, and the like, are developed in various types of pupils. Unless this is done we are unlikely to avoid creating "the organization man" in dependent-prone middle class students and alienated, social outcasts in lower class students.

Modern society requires continuous learning throughout life. If teachers do not promote adequate learning and the reassurance of progress, the student will come to reject learning - both in the school and in later life. Have we, as science educators, been so imprisoned by inhibited syllabuses, unadventurous teaching, restricted examinations and the academic tradition that we have failed to meet the most fundamental needs of our students? Perhaps in studying person-environment interactions in classrooms, the foundations of what Glasser (1969) calls "Schools without Failure", might be laid.

APPENDIX I

DISCRIMINANT ANALYSIS DATA

GROUP DESCRIPTIONS			GROUP CENTROIDS		
Group	Frequency	Gain	Root 1	Root 2	Root 3
1	High	High	29.5	4.7	-2.7
2	High	Low	26.4	5.5	-4.0
3	Mid	High	28.4	2.4	-3.1
4	Mid	Low	25.9	5.7	-6.0
5	Low	High	24.8	1.6	-5.3
6	Low	Low	20.5	4.3	-2.7

SIGNIFICANCE LEVELS FOR DISCRIMINANT FUNCTIONS

Wilks Lambda = 0.116;	$F = 1.819$;	$p = 0.0000$
Root 1: 43.6% variance;	$\chi^2 = 79.9$;	$p = 0.00000$
Root 2: 24.3% variance;	$\chi^2 = 52.3$;	$p = 0.0032$
Root 3: 13.4% variance;	$\chi^2 = 36.1$;	$p = 0.0738$

FUNCTION DESCRIPTIONS (Correlation of Root vs Variables)

Function I: Achievement Syndrome

Expectation Success	0.84	Attitude	0.53
Scholastic Aptitude	0.61	Model Status	0.47
Intelligence (B)	0.43	Need Ach.	0.32

Function II: Social Orientation

Warm, sociable (A+)	0.31	Sensitive (I+)	0.33
Restless, Inattentive (D+)	-0.32	Self Sufficient (Q2+)	-0.36

Function III: Role Fulfillment or Social Maturity

Rejection Status	0.41.	Model Status	0.32
(+ve) Expansiveness	0.53	Ego Strength (C+)	0.32

APPENDIX II

ANOVA RESULTS - RESIDUAL SCORES

ANALYSIS 1 : A = Class B = Frequency C = Anxiety

Significant F values were obtained for:

A Knowledge (p = .04) Comprehension (p = 0.4) Application (p = .00)
Total (p = .00)

BC Application (p = .13)

BC cell means were:

	High Anxiety	Low Anxiety
High frequency	0.4	2.0
Mid frequency	0.9	0.7
Low frequency	0.8	-0.1

ANALYSIS 2 : A = Class B = Frequency C = Anxiety

Significant F values were obtained for:

A Application (p = .00) Total (p = .02)

C Knowledge (p = .03) Comprehension (p = .00) Application (p = .02)
Total (p = .00)

BC Knowledge (p = .04)

AC Comprehension (p = .01)

BC cell means were:

	High ability	Mid ability	Low ability
High frequency	5.2	4.2	3.1
Mid frequency	5.3	3.8	4.1
Low frequency	3.4	6.2	2.0

ANALYSIS 3: A = Class B = Frequency C = Extroversion

Significant F values were obtained for:

A Knowledge (p = .00) Comprehension (p = .02) Application (p = .00)
Total (p = .00)

B Application (p = .06) Total (p = .05)

C Knowledge (p = .01) Comprehension (p = .00) Application (p = .04)
Total (p = .05)

REFERENCES

Cooley, W.W. and Lohnes, P.R. (1971). *Multivariate Data Analyses,* N.Y., John Wiley and Sons.

Gage, N.L. and Unruh, W.R. (1967). "Theoretical Formulations for Research on Teaching", *Review of Educational Research,* pp. 358-370.

Glasser, W. (1969). *Schools Without Failure,* Harper and Row.

Power, C.N. (1971). *The Effects of Communication Patterns on Student Sociometric Status, Attitudes, and Achievement in Science,* Unpublished Ph.D. Thesis, University of Queensland.

Siegel, L. (1967). "Integration and Reactions" in Siegel, L. (Ed.), *Instruction: Some Contemporary Viewpoints,* California, Chandler Publishing Co.

Siegel, L. and Siegel, L.C. (1967). "The Instructional Gestalt" in Siegel, L. (Ed.), op. cit.

Veldman, D.J. (1967). *Fortran Programming for the Behavioural Sciences,* N.Y., Holt, Rinehart, and Winston.

FORMATIVE EVALUATION AND THE AUSTRALIAN SCIENCE EDUCATION PROJECT

Gregor A. Ramsey

The Project is engaged in developing about forty discrete units of instructional material directed at junior secondary students in one of three identified stages of intellectual development and which conform to a content derived from the child's exploration of the environment. Brief statements of the Project's rationalé have been given in Newsletter No. 2 (ASEP, 1970) and more recently in the Australian Science Teachers Journal (Ramsey and Dale, 1971).

The major research task at this stage for the Project is the formative evaluation of the units produced. This evaluation is to be a continuing process throughout the total development of the units, with authors and evaluators having the closest possible opportunity for discussion and the inter-change of ideas. Also, to increase the validity of evaluative comment, each evaluator has drawn clearly to his notice the positions on science education being taken by the Project. To this end, a pamphlet 'Australian Science Education Project - Position Documents - A Summary' (ASEP, 1971) was prepared. In this document, the aims of the total package of materials is given along with statement on:

1 The main ideas to be developed in ASEP materials.
2 The criteria for the choice of topics for classroom study.
3 The ways of dealing with content, based on the stages of development of children.
4 The intended instructional strategies, for example, the use of the inquiry approach.

Trials teachers were selected only if they were in general accord with the positions taken by the Project and consulting evaluators are requested to evaluate the ASEP materials in the light of ASEP positions, rather than from the perspective of their own philosophy of science education.

The Project was conscious of attempting to ensure the three kinds of validity as outlined by Stake (1967) viz,

Content validity: Is the content as outlined in the materials good science?
Philosophical validity: Do the materials and strategies proposed fulfil the philosophical positions regarding the nature of knowledge and learning taken by the Project?
Pedagogical validity: Are the materials as presented, teachable with the target population of teachers and students and with the facilities that are available?

There were two categories of evaluative action which could be used in an attempt to maintain the validity of the materials. The first is reflective evaluation, where scientists in the field, science educators, sociologists, psychologists, and so on, read the materials, and express opinion on how well the units meet objectives as stated. The reflective evaluation of materials should to a large extent ensure content and philosophical validity.

The second category is active evaluation, where the materials in a form as close as can be afforded to their final quality are provided to a sample of the target population for evaluation in schools. Active evaluation is costly and time consuming, and in a Project with a fixed time limit, and schools which functionally operate for something less than 30 weeks per year this trial of materials presents serious practical problems. Even so, only from this active evaluation can information be gathered which will help the Project answer questions such as:

1 How long will the unit take in schools?
2 Which options do the students prefer?
3 What practical difficulties arise in the classroom?
4 What level of outcome can be expected from students?
5 Which method of presenting visuals is more effective?

In the evaluation of units, a balance must be struck between reflective and active evaluation and the amount of data that can be handled. Formative evaluation is a continuing process which should be designed so that when a unit emerges from the system, the developers can be reasonably certain that the unit will work with the target population, within certain limits that were established during the active trials.

The Project has a system of specification and sequential stages in the development of units designed to make certain that the units meet established criteria before they are made available for publication.

Step 1 **Broad Specifications**

The Area Specialists of Development branch decide a broad range of topics which may be developed as units within the Project's position statements. Each of these topics is then specified broadly in terms of the outcomes to be expected of the unit, and how the unit contributes to the aims of the total Project. More units are specified at this stage than can be completed within the time and resources available.

Step 2 **Checking Details of First Specifications**

From the total number of first specifications, those topics which on reflection seem most likely to foster the aims of the Project are selected for detailed specification. Also at this stage, evaluation of any proposed sequencing of units, decisions regarding the target population for the unit and likely necessary pre-requisites are determined.

Rejection of units at this stage is a relatively inexpensive process.

Step 3 **The Development of the Second Specification of a Unit**

The second specification is a detailed statement of intent for a unit based on the first specification, prepared by the team who are to be responsible for its development and formative evaluation.

The purposes of the second specification are:

1 To give a statement of intent for developing a unit which may be accepted, modified, or rejected before too much time is lost or before there has been too much personal involvement.

2 To act as a form of training for the developer, and to give the developer the necessary background information for preparing the unit.

3 To give comprehensive survey of the content, activities, and objectives of the unit so these may be appraised in relation to the position documents.

4 To explore choices among possibilities for the direction of development of the unit. The choices should be outlined in the specification and either pursued in parallel in the manuscript or rejected.

5 To give an opportunity for consultants and others interested to suggest desirable changes in the unit and to comment on the likely validity of the science content, the philosophy of the unit, and its teachability.

6 To give advance notice of any special film, apparatus, photo or other aids which may be needed so that development of these may be commenced while the unit is being written.

7 To give an opportunity for appraising the best methods of communicating to children the main ideas to be developed.

8 To propose a research design for the formative evaluation of the unit in terms of its stated objectives.

The second specification is written by a materials development officer (MDO), in conjunction with a discussant, a research officer and, in some instances, an outside consultant. This team, in general, will take the unit through to final publication stage.

In general, the MDO selects a unit from the first specifications for which he has the necessary background knowledge and interest to develop. He prepares the specification in consultation with the discussant and research officer. The discussant in most instances prepared the first specification, which forms the basis for the second specification. His role is advisory, on the main science ideas to be developed and on the practicability of the proposed activities.

The research officer's role during the writing of the second specification is

1 to remind the developers of any broad deficiencies in the first specification by relating it to the position documents

2 to help in the stating of the objectives for the unit in the second specification

3 to design the formative evaluation and trial of the unit

4 to co-ordinate the development of any tests or evaluation instruments to be used with the unit.

At this stage, the research officer must begin to resolve the dual nature of his role in the team. His first task is to make certain that sufficient feedback of the right kind will become available to the developer in any re-write of the unit. The second is to see that within the unit, sufficient means are given to the teacher and to the students to evaluate progress and learning outcomes. The instruments for the first area will largely disappear as the unit reaches final form. The instruments from the second will form part of the instructional package.

Step 4 Evaluation of the Second Specification

The second specification is a detailed statement of intent, and as such its evaluation should be broad with an idea to open up the unit to further possibilities which may be worth exploring.

Re-direction of the team's efforts or even rejection of the total unit at this stage means a much smaller wastage of work time than if such a decision was to wait until the whole unit is in manuscript form.

The second specification is circulated to Project staff and consultants outside the Project for detailed comment. A standard evaluation form is supplied which directs the reflective evaluation to the Project's position statements.

The purpose of the evaluation of the second specification is to give the development team suggestions for change, by attempting to answer the following questions:

1 How well does the specification of intent reflect the total aims of the Project, and the Project's philosophy for the choice of main ideas?

2 Is the science content as outlined scientifically acceptable?

3 Are the ideas proposed suitable for the stage of development of the students to whom it is directed?

4 Are there other more promising lines of development which could be pursued?

5 Are the student activities proposed workable in the classroom?

6 What new ideas relevant to the unit, interesting activities, or innovative tests could profitably be included?

7 Is the terminology used sound, and is there any looseness in the statements which may have to be watched in the final manuscript?

8 Are there any reference sources, or outside persons, helpful to the developer, not mentioned in the specification?

After evaluation information has been collated and presented in a form to help decide the future of the unit, the specification is made the subject of panel discussion with relevant Project staff and outside consultants.

Step 5 Writing of Manuscript

At this stage, the development team takes into account the advice from the second specification evaluation. The final form of the manuscript is planned and all activities are tried by the developed and one other person. Activities are tried with students where necessary .

During the period of writing the manuscript, the research officer plans the evaluation instruments. Discussions are held with the audio-visuals officer, laboratory technician and other service staff regarding specific requirements for the unit. All test materials, audio-visuals, and special equipment must be available at the same time as the printed materials are presented to trial classes.

Step 6 **Evaluation of Manuscript**

This stage is not intended to be a detailed evaluation. It is considered more important to have the unit produced in trial form and tried in schools than it is to reflectively evaluate it without the benefit of trials information.

The manuscript is approved by the development team, edited, and then passed to production where the materials are designed, diagrams and photographs prepared and the unit is printed in trial form. To the printed materials are added the audio-visual materials, special equipment and so on that make the complete unit package.

When the package is complete, the teachers selected for the trial of the unit are brought to the Project for a period of education on the philosophy of the unit, intended strategies to be used and methods of evaluation. These teachers are then ready to try the unit with their students, and they maintain weekly contact with the development team over the period of trial.

Step 7 **Evaluation of the First Trial Materials**

At this stage, evaluations from several sources are brought together to give directions for re-writing the materials, and for the wider national trial.

The main sources of reflective evaluation are the State Advisory Committees from each State, outside consultants, Project staff, and trials teachers. The active evaluation comes directly from the trial of the materials in schools. Pre- and post-test data, visits to trials classes by Project staff, trials teacher and student comments all form part of the data gathered from the active evaluation.

The purpose of the reflective evaluation is to ensure the content and philosophical validity of the materials. At the same time as the units are on trial, they are subjected to close scrutiny by scientists and science educators. The State Advisory Committees pay particular attention to any difficulties that may arise if the unit were to be used in their State.

The purpose of the trials is to test the pedagogical validity of the materials in the actual teaching situation. Schools close to ASEP headquarters were selected so that continuing contact could be maintained between trials teachers and Project staff. During trial, close attention is paid to errors, inconsistencies, and inadequacies in the materials as revealed by their use in the classroom.

In the first trial, a maximum of eight replications of each unit is undertaken. Each unit is tried in at least two high schools, two independent schools, and one of either a technical school or a catholic school. It will also be tried over a span of grades in an attempt to place the unit at its most acceptable level.

Suitable instruments e.g., questionnaires and checklists, are developed to facilitate the evaluation of trials materials in schools. A major source of such feedback will come from the fornightly meetings of trials teachers with Project staff.

Evidence gathered from the trials is used to help the development team answer the following questions:

1 How feasible are the learning experiences outlined in the units in different types of schools and with different teachers?
What problems arise when the materials are used?
Which instructional methods are successful; which unsuccessful?

2 What changes are necessary in the units to make them more useful?
How should learning experiences be restructured?
What new or additional activities are required?
Which activities are unsuitable?

3 What evaluation instruments are needed to help monitor?
How well students are achieving the objectives of the unit?
How teachers are using the unit?

4 How effective are the materials for communicating ideas?
What is their readability?
Can the students read them?
Do they want to read them?
How attractive are they?
How can we improve format, size, typeface, photographs, diagrams?
How well do they stimulate interest?
How could the interest in the unit be increased?
How does student interest compare with that for other units or other materials?

5 How does the teacher spend his time when using the materials?
How much time is spent preparing, organizing, maintaining discipline, giving verbal instructions, and so on?

Information to help answer these questions and to make appropriate decisions for change is gained from three major sources:

1 **Project staff:** The developers observe directly the materials being used in trial classes during their visits to schools, and have on the spot discussions with teachers and students. Ideas from all the teachers trialling a unit are discussed during the fortnightly meetings. Emphasis is placed on what students and teachers are doing and have to do in the classroom. Some audio and video tapes of lessons are made for discussion.

2 **Trials teachers:** Each teacher is provided with two copies of the teacher's guide, which contains copies of all student materials. On one copy of the teacher's guide, trials teachers write detailed comments for return to the Project. These form a basis for deciding modifications prior to national trial. The second copy may be kept at the school.

There are checklists and questionnaires to be completed, including records of time spent on various activities, problems of obtaining equipment, suggestions for useful references and audio-visual, class organization, and alternative ways of using the materials. The Project is very conscious of the need to reduce the amount of work required of trials teachers. To this end, the number of checklists and questionnaires is kept to a minimum.

3 **Students:** They are given questionnaires and checklists to find out what they thought of the unit. Also the results of any evaluation exercises are collated to give information on how the tests may be improved and how well the objectives are being achieved. Valuable information is obtained on students by examining their work books, and in interview. Some of the tests designed to help students include diagnostic, readiness, aptitude, basic skills, and level of intellectual development. These will become part of the final unit package.

The evaluation feedback from all sources is co-ordinated and collated into a form that is useful to the development team. The development team prepares a summary of feedback which is circulated to the panel responsible for the evaluation of the original second specification of the unit.

The members of the panel re-read the unit prior to an intensive evaluation meeting in which all aspects of re-writing, production and evaluation of the unit are discussed. From this meeting a set of recommendations for the re-development of the unit are prepared, and the summary of the evaluation and the recommendations are circulated within the Project and to State Advisory Committees.

Step 8 **Re-Writing of the Unit for National Trial**

Based on the recommendations from the summary evaluation, the development team re-writes the unit and it is produced for national trial. Again, the manuscript has only brief reflective evaluation before it is sent to all States for wider national trial.

At this stage the unit is again sent out to consultants for evaluating the content and philosophy of the materials in a manner similar to that described during the first trial.

Step 9 **The National Trial of Units**

The national trials of ASEP materials differ in many respects from trials undertaken by other projects. ASEP does not have a total package of materials nor a complete text to be tried in some predetermined order. It is not possible to evaluate a total program of ASEP and the total effect on children. Such evaluation must await the availability of a total package and a time scale of at least four years to make this possible. For the national trials the maximum number of units any one teacher can try with any one group of children is five. This represents only a small proportion of a total of (say) 35 units which may be used eventually to service a school curriculum.

The national trials of units are an essential part of formative evaluation. The information to be gathered is designed primarily to give evidence concerning changes to be made in the materials in the light of classroom trials in all States. Because of the relatively small number of units to be tried in any given class, only limited evidence may be gathered on the effects a total programme of ASEP materials may have on students and on teachers.

The validity of the final form of the materials and the extent of their eventual adoption in schools depends to a large degree on the outcomes of the national trials. The national trials are designed not only to ensure the validity of the materials for Australian schools, but also to help the States move towards the eventual introduction of the materials. Thus the formative evaluation at this stage moves from the evaluation of materials to the formative evaluation of a proposal for implementation.

The national trials will provide a nucleus of teachers experienced in ASEP philosophy and the use of its materials who will be available to help with the introduction of the published materials in the various States. They will also provide an opportunity for the different systems of education in the various States to evaluate the materials and in particular, their relevance to local conditions.

The specific purposes of the national trials are to expand the knowledge gained from first trials by:

1 determining the suitability of the various units in different classrooms

2 refining the content, structure and presentation of the units, based on evidence gained from the trials in the various States

3 exploring different combinations and sequences of units and their effects

4 finding out specific needs of the various States and modifying units where possible to account for these

5 establishing or confirming necessary pre-requisites for teachers and students using a particular unit

6 determining sources of equipment and aids for the units in the various States

7 developing checklists and other supporting materials to help the teachers more effectively introduce ASEP units into schools

8 providing a group of teachers experienced in the use of ASEP units

9 testing a model for the ultimate introduction of the published materials into schools.

The Project is convinced from its first trial experience that during formative evaluation, more effective feedback is obtained if the closest possible liaison exists between trials teachers and the Project, and the design of the national trials is such that no single teacher is expected to work alone.

The State Advisory Committees (SAC) in each State have appointed a State Trial Co-ordinator to co-ordinate the trial of all units in the State for the SAC and the Project. The trial co-ordinator is given time as part of his professional duties to work on Project matters.

Each unit of material is tried by one team in each State. Each team consists of a team leader who will co-ordinate and participate in the trial of each unit. The team leader is an experienced teacher who is also the co-ordinator of junior science in the school at which he is teaching. He is responsible to the trial co-ordinator for the trial of each unit.

Each team consists of

1 a team leader - a competent science co-ordinator

2 a fully-trained teacher (4 year trained) with a science degree and at least three years' teaching experience

3 a trained science teacher just out from teachers' college

4 a trained teacher who is teaching science, but who is not a fully-trained science teacher with a minimum of three years' teaching experience e.g., one who has less than four science units, or only two years' experience.

In selection, no special recognition is made of school systems except that the proportion of students in the various systems should be reflected in the proportion of teachers in all the teams in a given State.

It is impossible in the design of a trial of this nature to account for all the variables which may influence the outcome. Indeed, if experience from overseas projects is any indication, irrespective of school system, student ability, or State, the teacher is a most important determinant of the outcomes of the instructional process.

In the design of the national trials, some account is taken of the following variables in the trial of each unit:

1 differences among teachers
2 differences in ability and interest among students
3 differences among States
4 differences among class types

From trial of a total package of units some evidence will be gathered on how the following may affect the introduction of ASEP units in schools:

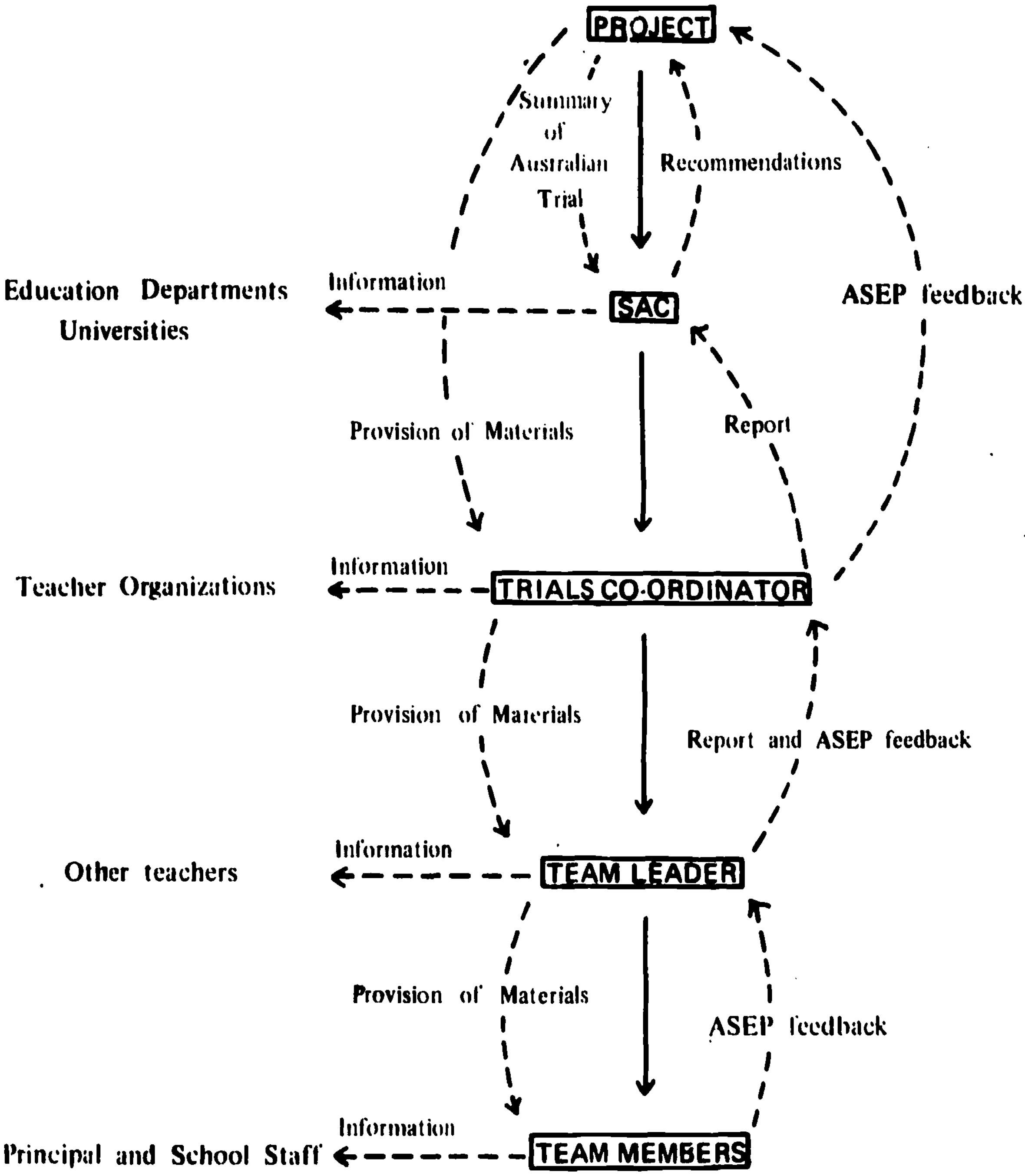
1 differences among school systems, e.g., State, independent Roman Catholic, and other independent schools, or country vs metropolitan
2 class size and composition
3 school facilities
4 existing State syllabuses
5 different sequences of units

For the trial of each unit, there will be two teams established in NSW and one each in the other five States. In the smaller States, the size of the team may be reduced. In addition four competent teachers as shown from their performance on the first trial of the unit in Victoria will try the unit in conjunction with the development team and so act as a control for the teams in the other States.

The feedback from within a State is collated before being returned to the Project. From the evidence gathered in the trial, a summary of recommendations for re-writing the unit is gathered. The unit is then re-written a final time and the special needs of the various States, for example, sources of supply of materials, are added. The materials are then made available for final production as the last step in the formative evaluation of the materials prior to publication.

The system is a long and time consuming process, and it takes up to twenty six months to put a unit through the total sequence with the resources available to the Project, although with a rate of up to one unit a fortnight emerging from the system, it takes less than three years to produce a total of 40 units. Whether the quality of materials obtained is worth the time and effort expended is still to be ascertained.

The following diagram represents lines of communication for the national trials.



REFERENCES

ASEP (1970). *Australian Science Education Project, Newsletter*, No. 2, Melbourne, ASEP.

Ramsey, G. and Dale, L. (1971). "The Australian Science Education Project", *Australian Science Teachers' Journal*, (in press).

ASEP (1971). *Australian Science Education Project. Position Documents - A Summary*, Melbourne, ASEP.

Stake, R. (1968). "The Countenance of Educational Evaluation", *Teachers' College Record, 68*, 523-540.

CHANGES IN PHYSICS STUDENTS IN THREE STATES OVER A TWO YEAR PERIOD

Lindsay D. Mackay

The Study

This report is based on an investigation of the changes that occurred for five samples of students during the two year period in which they studied five different senior secondary school physics courses.* The five courses concerned were the following:

(i) a course based on the Physical Science Study Committee (P.S.S.C.) physics course in Victoria in 1968 and 1969 (see V.U.S.E.B., 1967, 1968);

(ii) a pilot course based on the P.S.S.C. physics course in Queensland in 1968 and 1969 (see The University of Queensland, 1969);

(iii) the traditional physics course operating in Queensland in 1968 and 1969 (see The University of Queensland, 1967);

(iv) a pilot physics course operating in South Australia in 1969 and 1970 (see The Public Examinations Board of South Australia, 1969); and

(v) the traditional physics course operating in South Australia in 1969 and 1970 (see The University of Adelaide, 1968).

The samples in the five states involved in the study consisted of the following:

Victorian P.S.S.C. course	:	29 schools
Queensland P.S.S.C. course	:	8 schools
Queensland traditional course	:	6 schools
South Australian Pilot course	:	14 schools
South Australian Old course	:	14 schools

The ways in which these samples were obtained is described elsewhere (Mackay, 1971).

It was possible to compare the students in three of the samples with the population of students from which the sample was drawn. The marks obtained on the 1969 Matriculation Physics Examination by students in the Victorian P.S.S.C. sample were not significantly different from the marks obtained by the population of students presenting for this examination, and in terms of the results obtained on the 1969 Queensland Senior Examination in Physics, both the Queensland P.S.S.C. sample and the Queensland traditional course sample did not significantly differ from the relevant population of students presenting for the examination.

Students in the schools in the five samples were tested on three occasions:

at the beginning of grade 11 physics studies;
at the end of grade 11 physics studies; and,
at the end of grade 12 physics studies.

* This investigation was supported by a grant from the Australian Research Grants Committee in the years 1967-1971.

The test battery used on these three occasions was designed to test a broad range of possible educational outcomes that occur during senior secondary school studies. The objectives tested by this test battery are described elsewhere (Mackay, 1970), and will be briefly outlined with the results obtained later. As the sampling was conducted using schools rather than students as the basic sampling unit, for each school the mean scores on each testing were calculated for students who had been tested on all three occasions. These mean school scores on each scale on the three occasions were used as the raw data for subsequent analysis.

The five samples of physics students would not necessarily be expected to be equivalent at the beginning of senior secondary school physics studies for a number of reasons. Among these reasons is the fact that different proportions of students study physics in the different states (see Stranks, 1969), and the fact that different sampling methods were employed for the five samples. As the equivalence of the three samples at the beginning of the investigation could not be assumed, it was decided that in all analyses of the results an attempt would be made to control for differences in the initial aptitude of the samples.

Differences between the five physics courses in the changes that occurred in students during the two-year period of study of physics were tested using an analysis based on Winer's analysis of variance for a three-factor experiment with repeated measures on one of the three factors (see Winer, 1962, p. 337). The three factors in this analysis were:

Syllabus (5 levels);
Initial aptitude (2 levels based on scores on the Test of Cognitive Processes); and
Testing (3 levels).

Results

The results obtained on eight scales included in the test battery are represented in figures 1 - 8, and significant results obtained from the analysis of variance conducted for each scale are discussed briefly below. Note that only three measurements were obtained on each scale, and the mean scores on these three occasions are plotted in the graphs. These three points have been joined in the figures for convenience of representation.

Scale 1. Views About Physics Learning: (figure 1)

Objective tested: The student should view the process of physics learning as a non-authoritarian situation in which students are stimulated to think about physical phenomena, encouraged to "discover" physical relationships for themselves, and to participate in the development of experimental and theoretical methods for solving problems.

Results of analysis: There was a significant overall increase in scores on this scale during the two year period.

There was a significant difference between the five courses in the changes on this scale over this period (i.e., a significant interaction of syllabus by testing).

Scale 2. Views About Physics as a Process: (figure 2)

Objective tested: The student should come to view physics as an open rather than a closed process, which is by its nature dynamic, creative, tentative, and unfinished.

Results of analysis: There was a significant overall increase in scores during the two year period.

There was a significant difference between the five courses in the changes on this scale over this period. There was a significant difference between the changes for the "old" and "new" physics courses, but in the opposite direction to that anticipated originally.

Scale 3. Views About Scientists: (figure 3)

Objective tested: The student should view scientists as normal, active, occasionally fallible human beings who are different only in the area of their special training.

Results of analysis: There was a significant overall loss in scores on this scale during the two year period.

There was no significant difference between the changes for the five physics courses.

Scale 4. Enjoyment of Physics: (figure 4)

Objective tested: The student should come to view physics as an important activity for himself; he becomes committed to actively searching for an understanding of physical phenomena, and gains enjoyment thereby.

Results of analysis: There was a significant net decrease in score on this scale during the period in which senior physics was studied.

There was a significant difference in the changes on this scale for the five different physics courses.

Scale 5. Cognitive Preference for Theoretical as Compared to Applied Presentations: (figure 5)

Objective tested: Students should develop a greater cognitive preference for critical questioning and theoretical extension of information over practical applications of physical information.

Results of analysis: There was a significant overall increase in preference for theoretical as compared to applied presentations of physics material, but no significant difference in the changes for the different physics courses.

When "old" and "new" courses are compared, there was a significant interaction of type of physics course by initial aptitude by testing. From the profiles of this three-factor interaction, which are plotted in figure 9, it appears that students of high initial aptitude studying "old" physics courses shift towards a greater cognitive preference for applied as compared to theoretical presentations of physics material, whereas for the other three groups the shift is towards cognitive preference for theoretical presentations.

Figure 1

Mean Scores On Scale 1

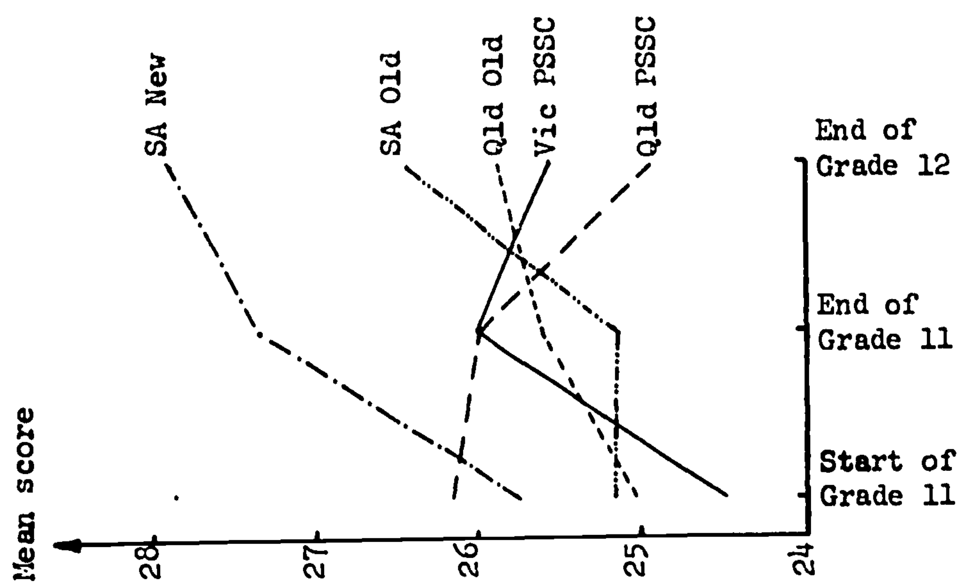
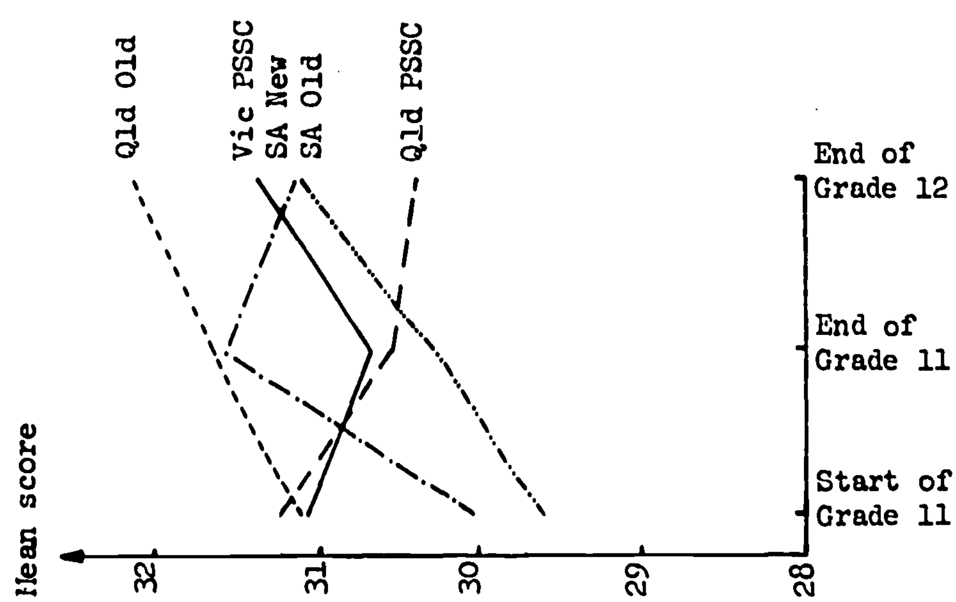


Figure 2

Mean Scores on Scale 2



112

Figure 3

Mean Scores On Scale 3

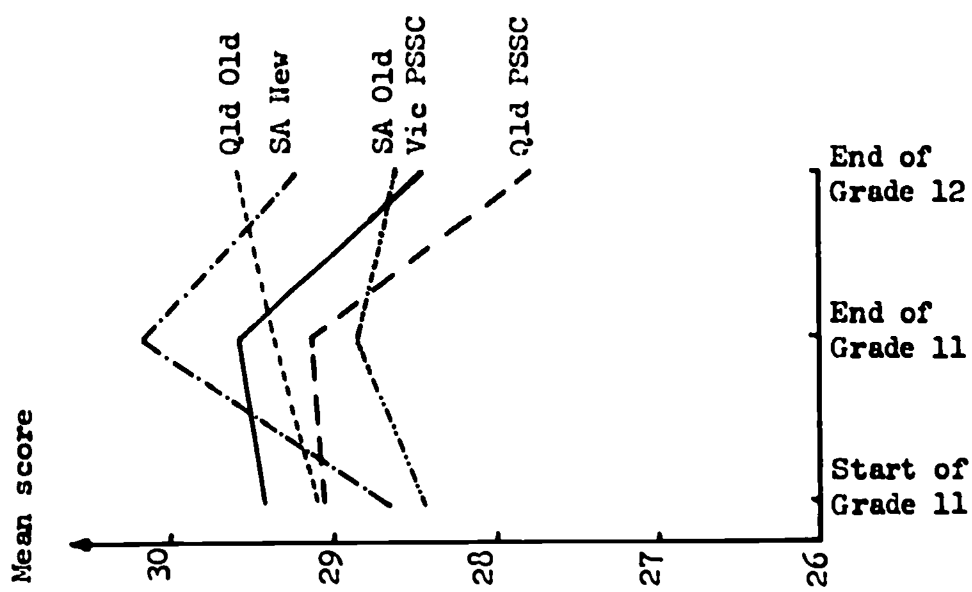
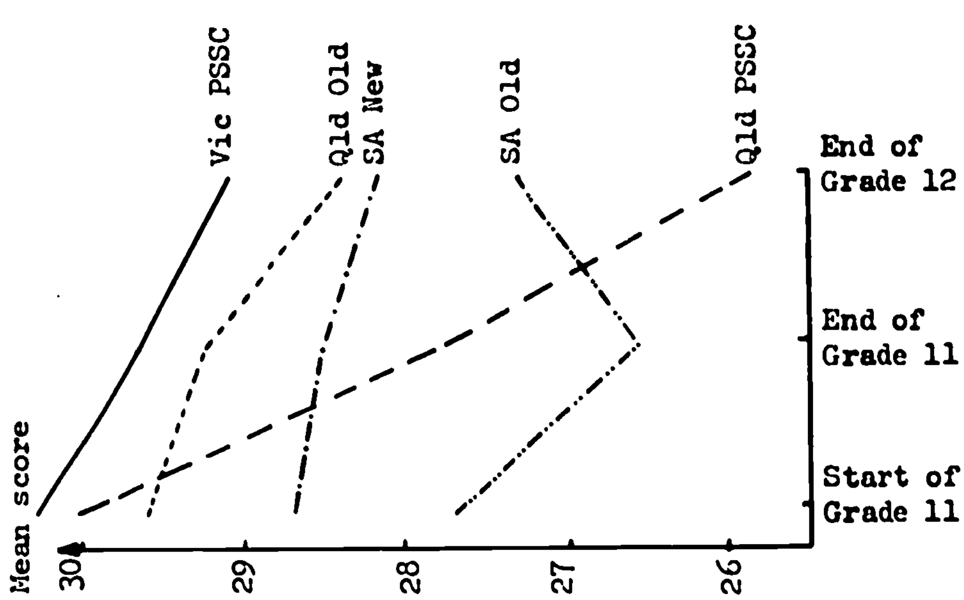


Figure 4

Mean Scores On Scale 4



Scale 6. Cognitive Preference for Understanding in Terms of Underlying Principles of Physics as Compared to Rote Memory of Specific Facts and Terms: (figure 6)

Objective tested: Students should develop a greater cognitive preference for identification of fundamental principles over rote remembering of specific fact and terms.

Results of analysis: There was a significant overall increase in score on this scale, but no significant difference in changes for the five courses.

Scale 7. Understanding of the Nature of Physics: (figure 7)

Objective tested: Students should come to understand the nature, function, and limitations of physical laws, to appreciate the purpose and features of models in science, to understand the relations between observations, theory, and experiment, and to understand the notions of assumptions, idealizations, and approximations.

Results of analysis: There was a significant overall increase in mean score on this scale, but no significant difference in changes for the five courses.

Scale 8. Cognitive Processes: (figure 8)

Objective tested: Students should develop ability to analyze problem situations mathematically.
Students should develop ability to use graphical presentations of data.
Students should develop ability to formulate simple scientific models.
Students should develop ability to make logical predictions based on a model.
Students should develop ability to make relevant observations.
Students should develop ability to suggest new lines of investigation based on observations.
Students should develop ability to make approximations and draw valid conclusions from observations and data.

Results of analysis: There was a significant overall increase in score on this scale, but no significant difference in changes for the five courses.

Figure 5

Mean Scores On Scale 5

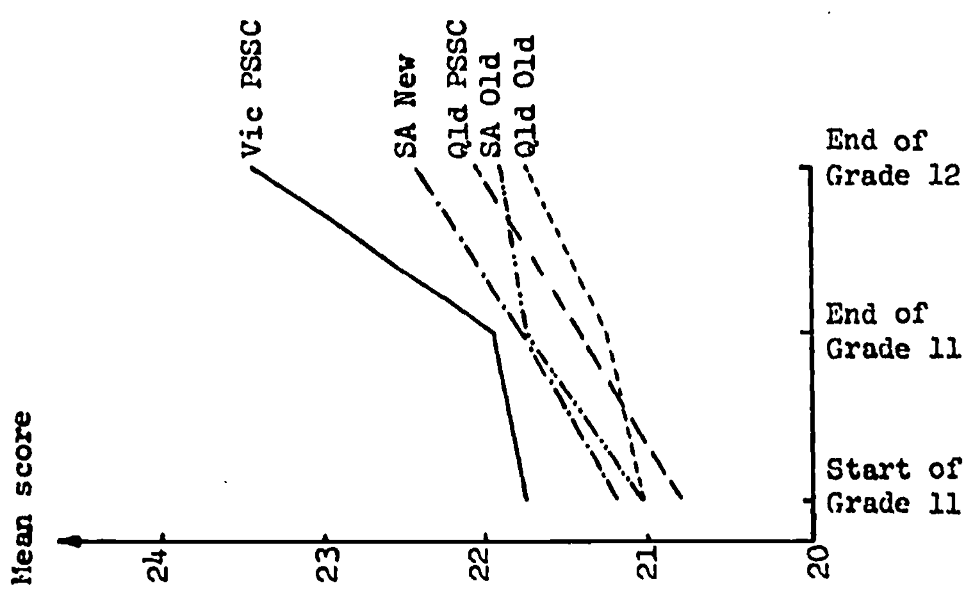


Figure 6

Mean Scores On Scale 6

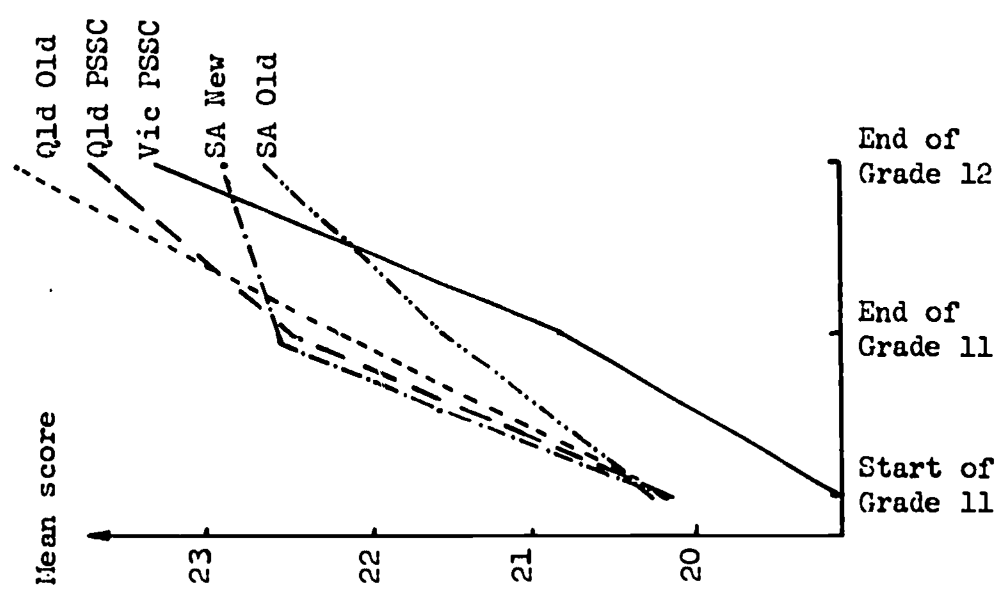


Figure 7

Mean Scores On Scale 7

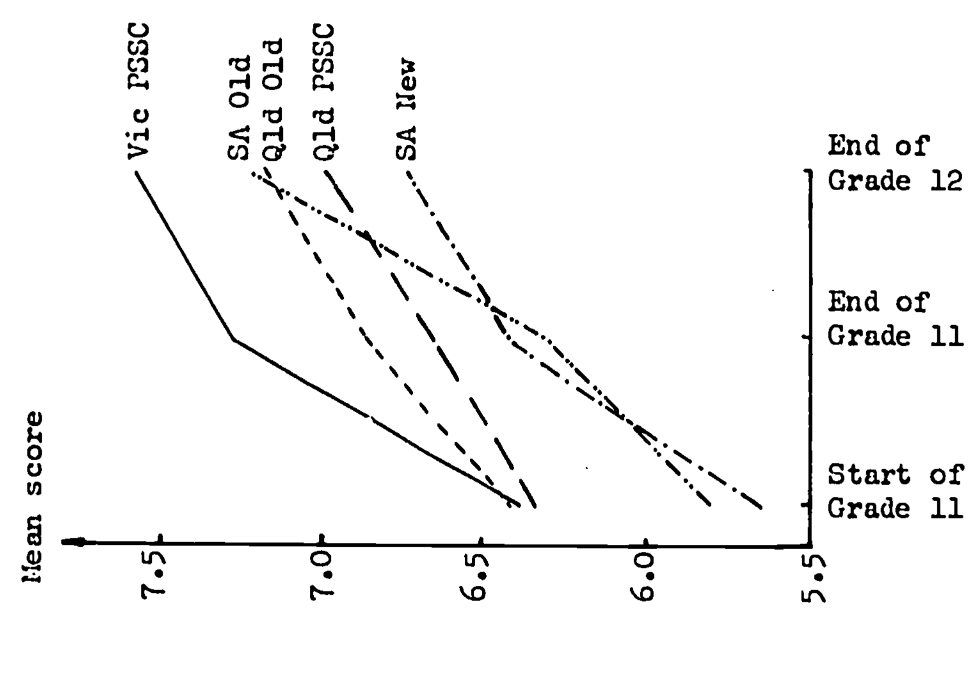
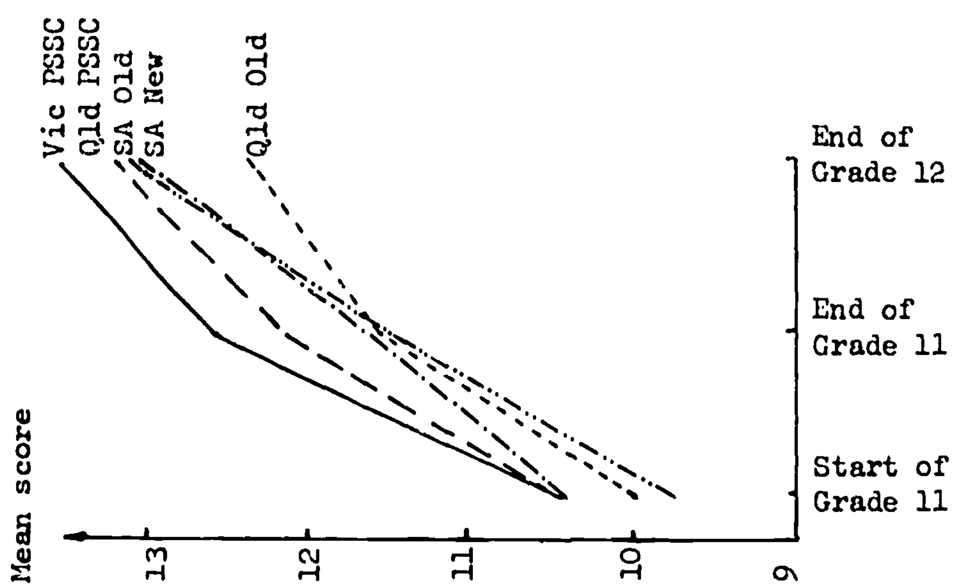


Figure 8

Mean Scores On Scale 8



Conclusion

The results reported above are of necessity an extremely brief summary of the results obtained in this study. One striking aspect of the results obtained is the similarity in the changes observed for students studying the five different physics courses. This similarity of the changes suggests that some of the stated differences between the "new" and the "old" physics courses may be imagined rather than actual.

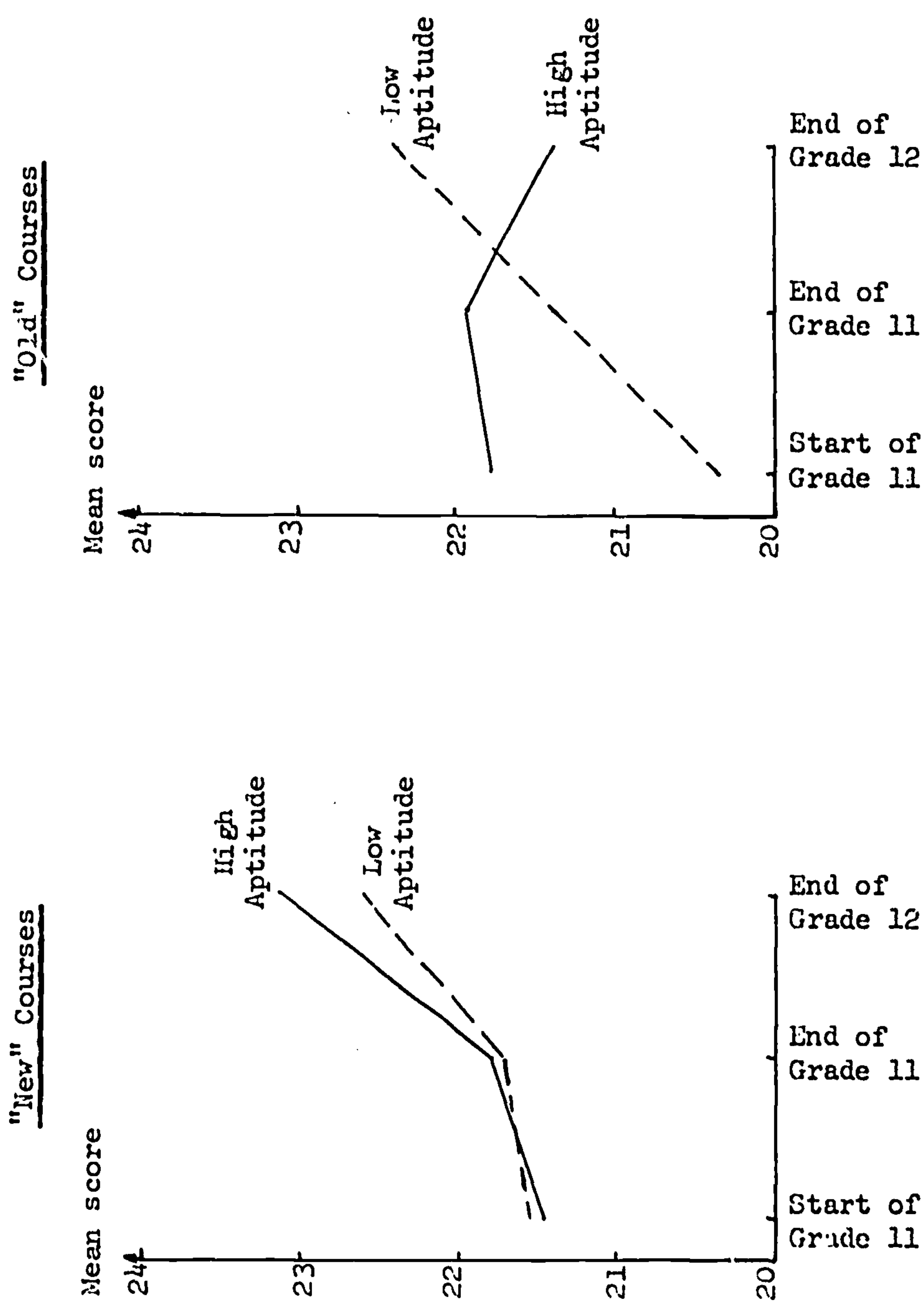
A number of points must be born in mind in interpreting the results reported in this paper. Firstly, the research design used facilitated the study of the changes that occurred in students during the period in which they studied senior secondary school physics, but these changes cannot be attributed to the study of physics. Secondly, other results indicate that there are as many different variations on a physics course as there are teachers, so that the validity of talking about the same physics course in different schools can be questioned. Thirdly, it is impossible to estimate the effects on student changes of the novelty or disruptive effects associated with the Queensland P.S.S.C. course and the South Australian pilot physics course operating as courses for the first time. Fourthly, the test battery used in this study tested only a limited range of possible outcomes of physics courses.

The research reported in this paper suffers from all the defects associated with studies involving the study of naturally-occurring educational settings.

Further research in this area is needed to investigate the changes in students associated with particular learning experiences in physics.

Figure 9

Profiles of the Three-Factor Interaction for Scale 5



REFERENCES

Mackay, L.D. (1970). "Changes in Victorian Physics Students during Two Years of Physics Study", *The Australian Physicist*, pp. 103-109.

Mackay, L.D. (1971). *Cognitive Preferences and Achievement in Physics*, unpublished Ph.D. Thesis.

The Public Examinations Board of South Australia (1969). *Manual of Syllabuses: Leaving and Matriculation Examinations 1970*, Adelaide.

Stranks, D.R. (1969). *Drifts from the Sciences*, Stanhope Oration delivered in Adelaide at CONASTA XVIII, September 1969.

The University of Adelaide (1968). *Manual of Public Examinations Board 1969*, Adelaide.

The University of Queensland (1967). *Matriculation Manual Senior and Supplementary Matriculation Examinations for the Years 1967-1968*, Brisbane, University of Queensland Press.

The University of Queensland (1969). *Matriculation Manual Senior Examinations for the Years 1969-1970*, Brisbane, University of Queensland Press.

V.U.S.E.B. (1967). *Handbook of Directions and Prescriptions 1968*, Melbourne.

V.U.S.E.B. (1968). *Handbook of Directions and Prescriptions 1969*, Melbourne.

Winer, B.J. (1962). *Statistical Principles in Experimental Design*, New York, McGraw Hill.

PROBLEMS IN THE SUMMATIVE EVALUATION OF EXPERIMENTAL SCIENCE CURRICULA IN DEVELOPING COUNTRIES

G.R. Meyer

Introduction

Developing countries, have been over the past decade, profoundly influenced by the new science courses emphasizing discovery oriented learning produced by curriculum groups in Great Britain and the United States. The various elementary and secondary science programmes sponsored directly, or indirectly, by the National Science Foundation (N.S.F.) in the United States and by the Nuffield Foundation in Britain, have been especially influential. Some have been adopted, virtually unchanged, by school systems in countries of Africa, Asia and South America. Others have been extensively modified by local educators to produce adaptations more accurately reflecting local educational objectives. In a few cases curricula have been developed from first principles by local groups, usually with the assistance of curriculum consultants provided under bilateral or international aid agreements.

Even in these latter cases, however, the influence of the new discovery oriented approach has been profound. It is possible to trace complex interlocking genealogies for almost all recent science courses in developing countries linking back to various N.S.F. or Nuffield projects.

In more recent years the time for intensive development and formative evaluation of those new courses has passed. Educators and politicians, and especially politicians, are now asking penetrating questions about the suitability and effectiveness of these programmes.

This paper discusses some of the problems involved and describes simple techniques used to obtain approximate assessments of the effectiveness and suitability of new science courses.

Key Problems of Curriculum Evaluation

Apart from the usual problems facing the curriculum evaluator in so-called "advanced" countries, notably problems of experimental design, the influence of the Hawthorne effect, or administrative or financial problems arising from working with an unsympathetic establishment, there are unique, or especially aggravated problems in developing countries.

The Origin of the Curriculum to be Evaluated

Vested interests are always influential in any country and in any school system experimenting with new curricula. In developing countries, however, this problem is frequently aggravated by the foreign origins of materials under review. The problem is even more difficult in cases where expatriate representatives of overseas curriculum projects are present as "field officers" directly involved in the trialling and formative evaluation of local versions of projects from their home countries. While in most cases these people genuinely strive for objectivity, some bias is inevitable. It is essential, therefore, that summative evaluation be undertaken by an impartial authority not previously involved in the development of the programme concerned. Unfortunately, this happens only rarely and in most cases "Caesar appeals to Caesar" and vested interest determines the outcome. The summative curriculum evaluator brought in from outside can be more objective, but must also be highly sensitive to the problems of the local field representatives of overseas curriculum projects.

The Design of Trials

As the summative evaluator is usually consulted late in the development of a new programme he has no opportunity to influence the pattern of trialling, the techniques used to obtain formative feed-back, or the gradual development of res· .rces available for evaluation. He must accept the situation as it is, and develop appropriate techniques to obtain as much relevant data as possible. In some cases highly organized representative experimental and control samples are involved with a clear experimental design and with a defined date for a conclusion to the experiment. In most cases, however, trialling is more informal, with the experiment diffusing to unsupervised schools and gradually, without proper evaluation, even to the whole school system. The samples may vary from three, or four carefully selected and closely supervised schools to dozens or even hundreds, often randomly selected and located in remote and inaccessable provinces. Supervision, in-service training and the extent and reliability of feed-back vary enormously from situation to situation. Methods used by summative evaluators must, therefore, be adaptable and flexible; subject to ready modification in response to local circumstances.

Shortage of Resources

Summative evaluation, if undertaken with academic vigour with very high levels of validity and reliability, is time consuming and expensive. Ideally, if high levels of accuracy are required, longitudinal studies over many months or years are necessary involving the development and use of refined instruments. Unfortunately careful long-term research studies are usually impracticable in developing countries for two reasons.

The first reason is political. Developing countries are in a hurry. They are rapidly changing, socially and economically, and educational innovation cannot wait for highly accurate research data, it must make do with approximations. Politicians must frequently make "snap" decisions and require "snap" data.

The second reason is the lack of resources. In the first place there is very little money available to develop research instruments, to employ staff or to analyze masses of data. More critically there is a shortage of trained personnel. Few employees of ministries of education, either local or expatriate, are trained even in general research methods, and almost none is trained in the specialized methodology of curriculum evaluation. There are few institutes or curriculum centres capable of diverting massive resources from the prime task of developing new courses or new learning materials to the task of long-term sophisticated studies of the effectiveness or suitability of these new courses of materials.

The curriculum evaluator in a developing country must therefore cut his cloth according to the material at hand. Methods must be simple, rapid and direct, and must be capable of application by personel with minimum training or experience. Cross-sectional, rather than longitudinal methods must be used, and data must be simple and easily analyzed and results must be easily understood by those not trained in the methods of educational research.

Political Influence

No innovation in education is free of political influence, but in developing countries at times these influences may be critical. There are problems at local, national and international levels. Politicians responsible for the introduction of trials into local or national school systems have a vested interest to see that their experiments succeed. At the national level changes in government or power struggles within a government, can profoundly influence attitudes and policies. Politicians naturally prefer evaluation studies supporting stated policies and the evaluator may be under some pressure to bias his findings accordingly. At a much higher level, the degree of clarity and precision in the statement of national objectives is of vital importance to the curriculum evaluator in determining criteria of suitability. In Tanzania for example, President Nyerere's "Education for Self Reliance" (Nyerere, 1967) gives a clear manifesto to the curriculum worker and the summative evaluator has a ready-made set of criteria by which to judge relevance and suitability. In other countries national objectives are less clearly defined and the problems of relevance and suitability become more difficult.

At the international level political influences may be complex and central importance. Many governments in ex-colonial territories are suspicious of any development that in any way seems to perpetuate the power and influence of the ex-colonial power. The curriculum evaluator must be seen to be uninfluenced by international politics and impartial in his judgement.

Because of these influences the techniques used by the curriculum evaluator must be as objective as possible and his report must be apolitical.

What Should be Evaluated?

The summative evaluator must answer central questions of significance to politicians and educators. Both groups wish to know whether to:

(a) abandon a new course and replace it by a new project;
(b) retain the course with modification after further trialling, or
(c) introduce the new course to all schools without further trial.

The politician usually has three questions. He wishes to know first if the new curriculum is relevant to national aspirations and consistent with national objectives. Second he would like an assessment of its practicability in terms of resources and finances. Third he is interested in whether the course is popular and well liked by pupils and teachers.

Educators ask additional questions about issues such as the degree of success in developing and understanding key concepts; success in attaining stated educational objectives; effectiveness and practicability of recommended teaching strategies; general levels of difficulty, and the nature of administrative problems associated with implementation.

The ultimate question that must take all these factors into consideration is "What is the probability of 'success' of this new course when transferred from trial schools to all schools?" The answer to this question determines policy.

After review of political, social and educational needs in several developing countries, the author recommends the following criteria by which to judge the "success" of a new course.

1 Achievement of basic concepts by the pupils
2 The acceptance and effectiveness of the recommended teaching strategies
3 Achievement of objectives in typical lessons
4 Relevance of content to life after school
5 Relevance of objectives to national aims
6 Extent to which administrative problems of the course have been overcome:
   (a) supply and maintenance of equipment and other resources
   (b) extent to which equipment has been improvised
   (c) suitability of the teachers.

Sum of criteria 1 to 6 gives the probability of success of the new course when transferred from trial schools to all schools.

An index of effectiveness can be estimated for each of the above criteria using a rating scale as follows:

Scale	0 - 20	Generally failed in effectiveness. Remedial action possible but difficult.
	21 - 40	Not successful but remedial action possible without too much difficulty.
	41 - 60	Acceptable level of effectiveness. Remedial action obvious and straight forward.
	61 - 80	Very satisfactory. Only minor problems which would be simple to overcome.
	81 - 100	Very high level of effectiveness. No remedial action necessary.

The art of the summative evaluator is obtain enough systematic evidance to enable these ratings to be made as objectively as possible. This evidence should be collected by a small panel of curriculum evaluators working intensively over a short period of say three to four weeks. All data obtained should be scanned somewhat impressionistically by all members of the panel who should develop independent ratings for each criterion, but where ratings differ compromise ratings should be determined after discussion.

The rest of this paper reports examples of these procedures developed and applied by the author in East and Central African countries in recent years.

The Nature Of The Evidence - Illustrated by a Case Study from Malawi

Evidence Required - The Methodology

Keeping in mind the limitations of resources and the general questions to be answered in an effective summative evaluation of a new curriculum, the following methods have been developed by the author in various African countries (Meyer, 1970, 1971). In each country trial schools associated with specific projects have been visited by a curriculum panel and at each school, the following methods applied.

a) A topic of central importance to the discipline likely to occur in first form programs in most school systems was identified. A short multiple choice objective test was prepared on this topic testing achievement of the cognitive objectives as defined by Bloom (1956). It was essential for administrative economy that the test could be answered by all pupils in only one period. It was scored immediately by the pupils who exchanged and corrected papers. In biology almost all African countries introduce the principles of biological classification during the first forms, so that topic was usually suitable for this purpose. The test was administered to Form I and all relevant higher forms to enable estimates of cognitive gain. No previous warning was given. It was administered and scored under the supervision of the curriculum panel and not by the class teacher.

b) Each pupil was asked to write a short statement on what he liked or disliked about the course. These statements were subsequently sorted into categories and scored by frequency of mention.

c) A specially prepared demonstration lesson by the classroom teacher was observed by at least two members of the curriculum panel. The teacher was invited to teach for maximum attainment of the cognitive, affective and psychomotor objectives of the course, and to use the recommended teaching strategies. The teachers' methods and the reactions of the pupils were observed, categorized and rated on various criteria.

d) One member of the panel presented a micro-lesson to a randomly selected class using the recommended teaching strategies - e.g., the discovery method by inquiry-oriented questionning. The reactions and skills of the class were observed, categorized and rated by the other member or members of the curriculum panel.

e) The school principal and the science teacher were interviewed about the acceptability and relevance of the new programme, and especially about specific problems and difficulties involved in the teaching. The interview was standardized and extensive notes recorded on a pro-forma.

f) The resources for teaching science in classrooms, library, laboratory and school grounds were observed, categorized and rated as poor, good or excellent by the curriculum panel.

After visiting an appropriate number of trial schools all data were systematically analyzed and tabulated. Each member of the panel then considered this data independently and in an overall impressionistic manner to determined ratings on the various criteria of effectiveness of the course. The panel then met to review ratings and to determine compromise ratings in cases of discrepancy.

The Case Study - A New Biology Curriculum in Malawi

From September 1967 to July 1968 UNESCO organized a workshop at Cape Coast, Ghana, attended by representatives of English-speaking African countries, to produce experimental editions of booklets and film loops for biology courses for grades 1 to 4 in African school systems. It was intended that each participating country would use these materials as the basis for locally developed curricula. The booklets were published in 1968 (UNESCO, 1968).

Malawi was represented at the workshop and a panel of biological educators from schools and universities was also established in Blantyre to review and implement the work of the international workshop at a local level. Beginning in 1969 certain units of the UNESCO course were tried in Forms 1 and 2 in five trial schools in Malawi and by April 1971 the Ministry of Education required a summative evaluation of the effectiveness of the new course as a course contributing to Malawi's Junior Certificate Examination given at the end of Form 2. The author was in Malawi in April and May 1971 as a UNESCO consultant to assist with this evaluation. The techniques and procedures outlined above were used and the results are reported below. Four of the five trial schools were visited, and some limited data were also obtained from two carefully matched schools following traditional programmes.

Results of the Evaluation Study in Malawi

The following results are given only to illustrate the application of techniques of curriculum evaluation developed by the author and described above in general terms.

Achievement Test: A standardized test of ten only multiple choice questions was administered to Forms1 and 2 in each of four trial schools and two control schools. The test assessed the following mental skills and biological concepts:

1 Knowledge that a key is useful for the identification of organisms.
2 Analysis of a situation involving the process of classifying.
3 Evaluation of the suitability of given words for naming a described plant.
4 Application of knowledge of plant characteristics to interpret growth under experimental conditions.
5 Evaluation of the most appropriate characteristic to give the least variable measurements.
6 Application of knowledge of plant and animal groups to the interpretation of a situation involving the classification of a group of organisms.
7 Evaluation of possible reasons for classifying organisms.
8 Comprehension of a pie chart in a specific case involving animal groups.
9 Evaluation of the suitability of given words to describe the features of described leaves.
10 Application involving the use of a biological key.

The frequencies of scores and the mean scores obtained by Forms 1 and 2 and the four trial schools and the two control schools are given in Table 1 and summarized in Table 2.

Results of this achievement test from the six individual schools are shown in Table 3. This table also summarizes certain characteristics of the schools.

The following generalizations emerged from these data -

1 While there was no statistically significant difference between mean scores of trial and control schools (Tables 1 and 2) there was a slight trend in favour of trial schools (see especially Table 3).
2 The test showed a good retention and understanding of material taught in Form 1 into Form 2.

3 While there was no overall statistical significance between performance of Forms 1 and 2, analysis of individual questions showed a definite trend towards deeper understanding of more difficult concepts in Form 2. (Questions 2, 3, 4, 5, 7, 9, and 10).

4 Comparing different schools, it appeared that whether the Headmaster was Malawian or not, whether the students were boys and girls or whether the laboratory facilities were poor or good, did not seem to significantly influence the achievement of pupils, which averaged about 55 per cent to 60 per cent in trial schools (note 40 per cent to 55 per cent in control schools). There was some suggestion however that pupils in boarding schools were more successful than those in day schools. It may also be significant that the only school with a rating for excellent for laboratory facilities was first in rank order of achievement.

TABLE 1

Frequency Distributions of Scores on Achievement Test in Four Trial Schools and Two Control Schools

Score	Trial Schools			Control Schools		
	Form I	Form II	Both Forms	Form I	Form II	Both Forms
10	0	2	2	0	0	0
9	2	5	7	0	0	0
8	9	15	24	1	3	4
7	21	15	36	5	9	14
6	33	37	70	13	11	24
5	27	36	63	15	14	29
4	27	16	43	15	9	24
3	7	5	12	14	5	19
2	2	1	3	4	1	5
1	0	0	0	2	0	2
0	0	0	0	0	0	0
N	128	132	260	69	52	121
Mean	5.4	5.9	5.7	4.5	5.3	4.8
Standard Deviation	1.46	1.53	1.51	1.51	1.43	1.55

Pupils' Attitudes and Interests: Pupils were invited to write freely on any aspects of the new course they liked or disliked. Their statements were categorized, scored by frequency of mention and scores expressed as percentages of numbers of pupils in the study. Separate analyses were made for trial schools and control schools. The results are given in Table 4.

TABLE 2

Mean Scores on Achievement Test in Four Trial and Two Control Schools

Forms	Schools Compared	N (Pupils)	Mean Score	Standard Deviation
I	Unesco trial	128	5.4	1.46
	Traditional	69	4.5	1.51
II	Unesco trial	132	5.9	1.53
	Traditional	52	5.3	1.43
I and II	Unesco trial	260	5.7	1.58
	Traditional	121	4.8	1.55

TABLE 3

Mean Achievement Scores in Four Trial Schools (1 to 4) and Two Control Schools (5 and 6)

School (In rank order of achievement)	Unesco Experiment or Control Traditional	Form I	Form II	Boarding or Day School	Biology Teacher: Nationality	Biology Teacher: Grad.	Headmaster	Boys or Girls	Quality of Lab. Facil.
1	Unesco	5.5	6.3	Boarding	Malawi	No	Dutch	Boys	Excel.
2	Unesco	5.9	5.6	Boarding	Malawi	No	British	Mixed	Good
3	Unesco	5.2	6.1	Day: self-help Boarding	Malawi	No	Malawi	Mixed	Poor
4	Unesco	5.5	5.5	Day	Malawi	Yes	Malawi	Mixed	Good
5	Control	4.8	5.2	Day	Malawi	No	Malawi	Mixed	Good
6	Control	4.0	5.4	Day	Malawi	No	British	Mixed	Good

126

TABLE 4

Percentage Frequency of Comments Made by Biology Students in UNESCO and Traditional Programmes

Aspects Liked	Form I		Form II		Form III	
	Unesco Experiment	Traditional	Unesco Experiment	Traditional	Unesco Experiment	Traditional
	N= 128	N= 69	N= 132	N= 52	N= 260	N= 121
1 Interest in plants and animals	73	83	59	58	55	72
2 Practical work	35	8	43	8	39	8
3 Ease of remembering and understanding	11	6	52	19	32	11
4 Relevance to life	12	20	26	12	19	16
5 Discovery approach. Independence of work and thought	15	1	24	0	19	1
6 Pupils books interesting and easy to read	6	0	19	0	12	0
7 Human body and Man's place in nature	9	30	16	10	10	22
8 Field work and outdoor collecting	4	13	5	2	4	8

Aspects not Liked	Form I		Form II		Forms I & II	
	Unesco Experiment	Traditional	Unesco Experiment	Traditional	Unesco Experiment	Traditional
	N= 128	N= 69	N= 132	N= 52	N= 260	N= 121
1 Some concepts difficult (especially terminology)	17	22	21	10	19	17
2 Some structures or organisms unpleasant or dangerous	11	16	11	0	11	9
3 Not enough equipment	1	0	7	0	5	0
4 Lack of success in performing experiments	1	1	9	2	5	2
5 Need for supplementary work, e.g. too many notes needed	7	0	3	4	5	2
6 Practical work	3	10	7	6	5	8
7 Field work and outdoor collecting	2	12	9	0	5	7
8 Pupil books difficult or lacking	3	0	5	4	4	2
9 Having to think for oneself	3	0	5	0	4	0
10 Not enough time for work	1	4	5	0	3	3
11 Some concepts uninteresting (e.g. plants or insects)	5	13	1	12	3	12

Results in Table 4 showed the following trends -

1 More Form 1 pupils mentioned their interest in plants and animals than Form 2 and this type of interest was generally higher in the control schools than the trial schools.
2 In almost all other categories liked there was a significant *improvement* from Form 1 to Form 2 in the trial schools. In the control schools there was either no change or a decline in interest from Form 1 to Form 2 in seven of the eight categories.
3 Almost 20 per cent of the pupils in the trial schools expressed interest in the "discovery" method but this was mentioned by only one pupil out of 121 in the control schools.
4 Pupils in control schools showed less interest in practical work than those in trial schools.
5 Interest in factual material such as the structure and identification of plants and animals or the structure and functions of the human body was greater in the control schools than the trial schools.
6 In Form 2 more than half of the pupils in the trial schools compared with 19 per cent in the control schools mentioned the ease of remembering and understanding the material.
7 The number of pupils commenting unfavourably on either trial or control courses was much less than those commenting unfavourably.
8 Criticisms of the courses were mainly in terms of the difficulty of selected concepts and in the terminology required and there was little difference in this regard between control and trial schools.
9 About 10 per cent of pupils in both courses commented unfavourably on unpleasant aspects such as dangerous organisms, noxious smells, sickening dissections or the need to kill harmless animals.
10 In general there was strong acceptance of the basic objectives of the programme and a fair appreciation of its relevance to Malawi and to life after school.

Observation of Lessons: One member of the curriculum team gave a fifteen minute lesson to a selected class in each of the four trial schools. Pupils were required to solve an unseen problem by observation and by questionning the teacher. Their responses were rated on a percentage scale by the person giving the lesson and by one observer. The results are given in Table 5.

TABLE 5

Achievement by UNESCO Trial Classes in Solving an Unseen Problem Presented as a Micro-Lesson. Consensus of Independent Percentage Ratings by Two Observers. (Percentage Ratings)

	Response	School (In Rank Order on Achievement Test - see Table IV)				Average Rating
		1	2	3	4	
1	Willingness to answer questions	90	80	20	80	68
2	Level of interest	90	60	50	80	70
3	Success in solving the problem	60	70	30	70	58

Each UNESCO teacher was asked some days before the visit to prepare a demonstration lesson. This lesson was presented at the time of the visit and observed by the curriculum workers. Ratings of the extent to which the lesson contributed to the objectives of UNESCO Biology were made on a ten point scale. The extent to which recommended testing methods were used was also noted and expressed on a ratio of desirable to undesirable (e.g., problem solving to verification).

The results are shown in Table 6.

TABLE 6

Achievement of Objectives of UNESCO Biology in Demonstration Lessons by UNESCO Trial Teachers (Concensus of Ratings [1 - 10]) by Two Observers

Objective	Schools (In Rank Order on Achievement Test) 1	2	3	4	Average Four Schools
A. Knowledge and Understanding					
1 Knowledge of facts in text	5	6	5	4	5
2 Understanding of facts in text	5	6	4	4	5
3 Response to questions	7	6	5	5	6
4 Understanding of principle of experiments	-	5	5	2	4
5 Understanding conclusions from experiments	-	5	4	2	4
6 Ability to reach independent conclusions	6	5	4	2	3
Average Knowledge and Understanding	6	6	5	3	5
B Emotional Reactions					
7 Interest during introduction to lesson	6	2	7	8	6
8 Interest during main part of lesson	6	3	7	6	6
9 Interest during conclusion of lesson	4	3	5	2	4
10 Interest in demonstration experiments	5	4	4	5	5
11 Interest in experiments performed by pupils	6	3	6	7	6
12 General attitude to biology lessons	5	4	5	8	6
13 General scientific attitude	6	4	4	4	5
Average Emotional Reactions	5	3	5	6	5
C Practical Skills					
14 Making biological drawings	-	-	4	3	4
15 Handling glassware	-	6	6	-	6
16 Handling dissection instruments	-	6	-	-	6
17 Handling biological specimens	6	5	7	-	5
18 Handling measuring instruments	-	-	-	-	-
19 Using a lens	-	4	-	-	4
20 Using a microscope	-	4	6	4	5
Average Practical Skills	6	5	6	4	5

continued on next page

Table 6 (continued)

Objective	Schools (In Rank Order on Achievement Test)				Average Four Schools
	1	2	3	4	
D. <u>Recommended Teaching Methods</u>					
21 Problem solving/verification	5	3	8	7	6
22 Pupil active/teacher active	3	2	4	3	3
23 Objects, materials/blackboard	5	4	8	7	6
24 Pupil experiments/ demonstrations	6	4	10	9	7
Average Teaching Methods	5	3	8	7	6

Summary: Average Ratings Four Schools -

Pupils Knowledge and Understanding	= 5	Total Achievement of Objectives 50%
Emotional Reactions	= 5	
Practical Skills	= 5	
Use of Recommended Teaching Methods	= 6	

Table 5 suggested that on the whole pupils in trial schools were interested in and responsive to problem solving situations. By Form 2 their ability to reason through an unseen problem was very satisfactory.

In Table 6 the results suggested that on the whole trial teachers were about 50 per cent effective in achieving suitable teaching objectives in specific lessons. In lessons observed the four teachers made use of recommended teaching strategies to varying degrees - 30, 50, 70, and 80 per cent of the time respectively.

Resources of Trial Schools: Table 7 summarizes the resources available in the four UNESCO trial schools. This table is based on interviews with headmasters and UNESCO Biology teachers and on a survey of the facilities at each school made at the time of the visit. Ratings were on a three point scale - excellent, good or poor and are a consensus of the opinions of the curriculum workers. There were few differences between the ratings of individual observers.

The following generalizations emerged from Table 7 -

1 Facilities did not seem to have had a strong influence on the course. Satisfactory standards and adequate achievement of objectives occurred in schools with excellent facilities and with relatively poor facilities. This was strong evidence in favour of the suitability of the course for the average secondary school in Malawi.

2 In spite of feelings of inadequacy in some teachers the course had been reasonably successful in attaining its objectives. There seemed to be little correlation between staff attitudes and the results. This might be taken as an argument that the course, perhaps throught the structure of its materials, could be taught successfully by teachers of widely different attitudes and temperaments.
3 All teachers under-estimated their success in implementing the objectives of the course.
4 None of the teachers had used much initiative in improving equipment or methods of teaching.

TABLE 7

Resources in Four UNESCO Trial Schools
(Rated by Two Judges as Excellent, Good or Poor)

Resource	School (In Rank Order on Achievement Test)			
	1	2	3	4
A. Technical				
1 Adequacy of school grounds	Excellent	Excellent	Good	Good
2 Basic apparatus such as glassware	Good	Good	Poor	Excellent
3 Specialised apparatus, e.g. microscopes	Good	Poor	Good	Good
4 Laboratory space	Excellent	Excellent	Poor	Good
5 Laboratory fittings - water, electricity, gas	Good	Good	Poor	Good
6 Equipment Innovations	Poor	Poor	Poor	Poor
7 Extracurricular enrichment	Good	Good	Poor	Poor
B. Staff (Unesco Teacher)				
1 Academic qualifications	Good	Good	Poor	Excellent
2 Attitude towards Unesco Biology	Good	Poor	Good	Good
3 Feeling of adequacy in teaching the Unesco course	Poor	Good	Good	Good
4 Success in implementing philosophy of Unesco Biology	Poor	Poor	Poor	Poor

Summary Evaluation: Table 8 gives ratings (on scale 0 per cent to 100 per cent) for a number of criteria of acceptability of UNESCO Biology in secondary schools in Malawi. It compares pupils and teachers. The evaluators drew somewhat impressionistically on all the data.

TABLE 8

Effectiveness of UNESCO Biology Forms 1 and 2 in Malawi
(Summary % Ratings by Curriculum Evaluators)

Criterion	Ratings for Teacher (Effectiveness as seen by teacher or teachers effectiveness as observed)	Ratings for Pupils (Effectiveness as shown by pupil achievement)
Achievement of basic concepts by the pupils	45	60
The discovery of approach		
a) its acceptance	60	80
b) its effectiveness	60	75
Achievement of objectives in typical lessons	50	60
Relevance of content to life after school	40	30
Relevance of objectives to national aims	85	30
Extent to which problems of the course have been overcome:		
a) supply and maintenance of equipment	85	90
b) extent to which equipment has been improvised	10	-
c) suitability of teachers	45	-
Chance of success if transferred to other schools	90	-

As a result of the evaluation study the curriculum Panel was able to report to the Ministry of Education in the following terms (Meyer, 1971) ·

> "The results of the evaluation study are most encouraging. An important finding is the satisfactory level in attainment of objectives of the course in spite of some unsatisfactory aspects of the teaching. Classroom techniques and feelings of inadequacy of the teachers are problems that can be easily overcome by in-service training and by improving the learning materials.
>
> Part of the general feeling of inadequacy and relatively low success in some areas can possibly be traced to a weakness in structure of the UNESCO booklets in their present form (the 1968 experimental edition). Teachers are unsure whether to treat them as conventional texts for reading; as workbooks or as combinations of these; and the approach in this regard varies from unit to unit. The teachers' guides in their present form do not help in this and are also deficient in other respects.
>
> A welcome finding has been the steady growth in understanding and acceptance of biological principles and of the philosophy of the course from form to form. Related to this are the high levels of interest and the feeling that it is relevant to pupils in later life, but in re-writing the materials an attempt should be made to make this more directly obvious to the pupils. Apart from this disparity between pupil and teacher reactions it is interesting that the teachers have generally underestimated their pupil achievements and their attitudes to the discovery method.
>
> The assessment by the evaluators that the course has a 90 per cent chance of success in average and even below average schools in Malawi, is of vital significance. It would appear that the problems revealed in the evaluation study are relatively minor and can be easily overcome. It is strongly recommended therefore that the Ministry of Education gives serious consideration to introducing this programme to all secondary schools, Form 1 and 2, in, say, 1973".

Conclusion

The simple techniques described in this paper can be applied rapidly and inexpensively by personnel previously untrained in techniques of curriculum evaluation. They are unrefined and open to much serious methodological criticism. Nevertheless they work. They are a way of giving those who must make value judgements about the implementation of new curricula some systematic information on which to base opinions. By using such methods for example, the author was able to show that the experimental course in biology in Malawi had a very high chance of success (90 per cent probability) when implemented in all schools in Malawi. Similar studies of current experimental programmes in Kenya and Zambia have provided interesting comparisons. In Kenya the Secondary Science Project in Biology was shown to have an 80 per cent chance of succeeding when spread from trial schools to all schools. In Zambia by contrast the Secondary School Science Pilot Project of the Science Education School of the University was shown to have only a 20 per cent chance of success. These indices provided valuable guidance to the Ministries concerned (Meyer, 1971). The use of such simple techniques has not been previously employed in developing countries and changes of policy have been based invariably on unsystematic opinion. The methods reported in this paper, therefore, appear justified in the absence of practicable alternatives.

REFERENCES

Bloom, B.S. (Ed.) (1956). "Taxonomy of Educational Objectives. The Classification of Educational Goals". *Handbook I Cognitive Domain,* New York, McKay.

Meyer, G.R. (1970). *Report to Division of Science Teaching, United Nations Educational Scientific and Cultural Organization on Three Months as Consultant to the Biology Panel of the Tanzania Institute of Education, University of Dar es Salaam on Aspects of Testing and Evaluation - SSP Biology, April to July 1970.* Paris, UNESCO.

Meyer, G.R. (1971). *Report to Division of Science Teaching, United Nations Educational Scientific and Cultural Organization of Eleven Weeks as Consultant to the Governments of Kenya, Tanzania, Zambia and Malawi on Aspects of Biology Teaching* - February to April 1971. Paris, UNESCO.

Nyerere, J.K. (1967). *Education for Self Reliance.* Dar es Salaam, The Government Printer.

UNESCO (1968). *Investigating Living Things. Booklets 1 to 12. UNESCO pilot project for biology teaching in Africa.* Experimental edition. Cape Coast, University Press Ghana.

"CAN SCIENTIFIC ATTITUDES BE EVALUATED?"

David Cohen

Introduction

One of the reactions to the perceived shortcomings of school science in USA following the 1957 launching of Sputnik I resulted in the involvement of scientists in curriculum teams (Cohen, 1964). The presence of these scientists caused a revival of interest in having school students mirror the scientist in the laboratory. Concurrently with the American space age trauma, the publication of the *Taxonomy of Educational Objectives* (Bloom, 1956 and Krathwohl, 1964) sensitized educators anew to the existence of two domains encompassing wide ranges of attitudinal and cognitive objectives. The *Taxonomies* highlighted the needs for more specific statements of objectives, and for a closer relationship between objectives, classroom experiences and evaluation procedures.

Previous Research

In science teaching, the development of scientific attitudes has been proclaimed as a key objective for at least half-a-century (Rowlands, 1971; Curtis, 1932). The effect has been a proliferation of statements on scientific attitudes, dating from the early 1920's, to the most recently published Australian statement (Australian Science Education Project, 1970).

Several attempts have been made to develop tests for evaluating scientific attitudes, e.g. Curtis (1924); Noll (1935); Cohen and White (1968); Rowlands (1971). Rarely have attempts been made to evaluate student progress towards these attitudes. *At the best,* the lists of scientific attitudes have constituted the materials which have often been cited in the forewords of science curricula.

However, although attitudes have been given lip-service, curriculum documents have given little guidance to teachers concerning appropriate related classroom-learning experiences. In an attempt to examine the development of scientific attitudes, Rowlands (1971) found that Victorian junior-secondary school pupils did not prefer natural over supernatural causes; have great faith in reason, discussion and experiment and prefer qualified generalizations to sweeping statements.

Plan of the Investigation

(i) Subjects and materials

In this present survey, it was assumed that scientific attitudes were those which would be agreed to by a group of scientists who had achieved eminence in science.

How does one determine who are the "eminent scientists" of Australia? The criterion which was adopted, was membership of the Australian Academy of Science. Membership of this Academy is open to "British subjects who at the time of their election are normally resident in Australia and are eminent by reason of their scientific attainments and their researches in natural science". Membership is based upon election by existing Fellows, following rigorously prescribed and exacting nomination, voting and selection procedures. Election is limited to six Fellows per year.

Of the total Academy membership in 1969 of 123, 114 were approached, and 64 were subsequently interviewed, individually. These scientists represent a "Who's Who" of Australia's most eminent scientists. However, it is appropriate to note that the selection and election procedures for membership tend to preclude, or delay membership for outstanding younger scientists, so that the "validating" group may be biased in terms of age. Also, during 1969-70, an extensive literature survey resulted in the location of 45 published sources of sets of statement on scientific attitudes. Upon elimination of replications and gross ambiguities, 230 distinct statements were retained. (A scientific attitude was regarded as an emotional disposition towards objects, ideas or activities related to science.)

(ii) Experimental Procedure

Sets of the 230 selected statements were printed onto cards numbered on the reverse side for ease of subsequent identification and coding of statement. The cards were shuffled for presentation to the scientists in their laboratories or offices. In keeping with the statistical techniques required for the Thurstone and Chave method of equal appearing intervals (Edwards, 1957) each scientist was asked to sort the cards into one of nine envelopes. The extremes of the range were marked on envelopes on which the numbers "1" to "9" were written respectively. These labelled envelopes were placed in front of the scientists. On Envelope No. 1 was printed, "Least essential components of the scientific attitude", and on Envelope No. 9 was printed "Most essential components of the scientific attitude".

To simplify the nature of the judgements required, a two-stage sorting process was used. The scientists were first asked to rate the "least essential" to "most essential" components across *three* intervals, and then to spread their three intervals "1", "2", and "3" respectively across new 1-2-3, 4-5-6, and 7-8-9 intervals. In each case only the two extreme intervals were defined for the scientists, as required by the Thurstone-Chave approach.

The category of the nine possible intervals into which each of the 230 cards was rated by each of the 58 scientists was recorded onto a coding sheet, and transferred to punched cards. Medians and interquartile ranges were computed for each of 230 statements.

Statement of Results

1 On only 30 of 230 statements of scientific attitude was there a high degree of consensus (i.e., $Q < 3$ on the 9-interval scale). Of these 30, 26 had scale values exceeding 7. These 30 statements are presented in Table 1.

TABLE 1

Statements of Components of Scientific Attitude on which Subjects showed a High Degree of Consensus (Q 3.0) and Highest Medians (S 6.0) (Ascending Q)

Statement No.	Components of Scientific Attitude	Interquartile Range (Q)	Median (S)
200	preparedness to revise opinions in the light of additional, reliable data	0.69	7.82
34	a creative imagination	1.24	8.20
32	a curiosity, a thirst for new knowledge and a desire to explore the unknown	1.31	8.09
228	intellectual honesty in communicating scientific findings; (the truth must not be suppressed)	1.43	8.10
13	honest and truthful	1.47	8.12
37	willingness to invent and use new methods of enquiry	1.48	7.55
154	willingness to consider novel hypotheses and explanations	1.79	7.76
175	preparedness to retest hypotheses against new evidence, modifying or rejecting them if necessary	2.96	7.85
107	preparedness to word clearly, precisely and unambiguously, the statement of a problem, an hypothesis, or a law	2.06	7.53
88	critical in thinking habits, prepared to use observations to question knowledge, and knowledge to question observations	2.14	7.36
44	perception of problems as a challenge and a determination to solve them	2.17	7.56
35	willingness to test intuition by deduction	2.19	6.66
185	recognition of the assumptions involved in the formulation of hypotheses and conclusions	2.39	7.47
97	determination to be careful, factual and accurate in all observations at all levels of enquiry, and evaluation of data	2.46	7.21
141	perception of relationships between important ideas or observations	2.48	7.33
149	distinction of a fact from an assumption	2.48	7.46

. . . . .

TABLE 1 cont.

134	disallows personal bias, pride, prejudice, ambition or wishful thinking to pervert the truth	2.60	8.77
174	willingness to recheck for possible errors of interpretation at all levels of enquiry	2.60	7.14
207	willingness to use every conceivable means to prevent, to discover and to correct possible errors in your work	2.64	7.50
206	preparedness to use control experiments	2.66	7.31
224	preparedness to admit being wrong	2.70	7.90
198	preparedness to verify conclusion by repetition and comparison of experiments	2.72	7.24
26	liking for initiating and carrying out investigations	2.79	7.52
177	recognition of which factors must be controlled in an experiment	2.81	7.10
158	awareness of and clear statements of all assumptions	2.84	7.42
133	preference for reliability and accuracy in procedure, calculation, observation and the reporting of these	2.88	7.03
108	shuffling of facts in the mind, rearranging, reorganizing and classification in order to define the problem	2.89	6.10
176	willingness to gather additional data if necessary, in order to test the hypothesis	2.90	6.87
28	sense of enquiry	2.91	7.04
139	care, criticism and accuracy in the use of data collected by others	2.95	6.58

2 Attitudes on which scientists agreed, but for which their *rating* (scale value) was *low*, included consideration of possible future developments from or consequences of their work.

3 Attitudes upon which scientists lacked consensus included tolerance of outlook, the question of orderliness in the universe and in nature, recognition of the limitations of science and of their fields of authority.

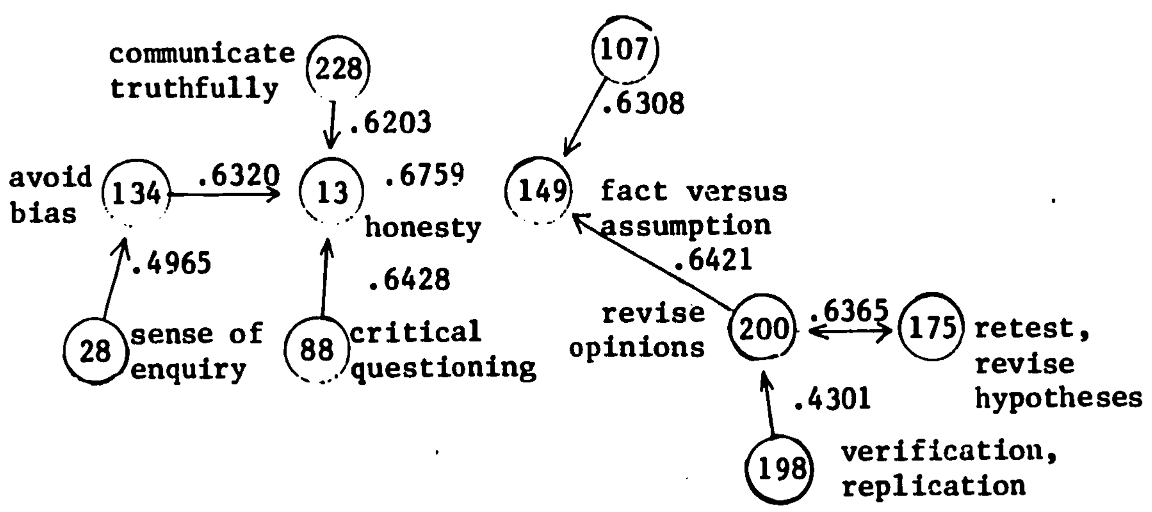
4 As many scientists considered that beliefs in "superstition, sixth-sense, charms, omens, good or bad luck, and appeals to the supernatural" was an essential component of their scientific attitudes, as those who rejected these beliefs.

138

5 Honesty, truth, overcoming bias and prejudice, and the avoidance, admission and exposure of errors were regarded by most scientists as essential components of scientific attitudes.

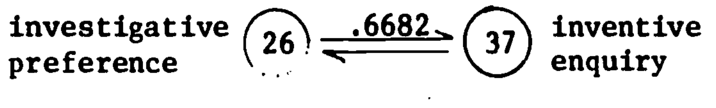
A product-moment correlation matrix was produced to explore the extent of relationship of scientists' ratings between the thirty high consensus, high scale value statements. A large number of the correlations were high. A technique was used to identify "sets" of statements. Clusters were formed using a simplified form of McQuitty's Cluster Analysis technique. This method was preferred to factor analysis, since it eliminated the need to make assumptions about the shape of the distribution and overcame possible problems of matrix inversion. Visual examination of the 30 x 30 matrix yielded pairs of items with correlations which were highest with each other (i.e., reciprocal pairs). By working through the highest inter-correlations, clusters are built up, and interpreted. The six clusters are reproduced below:

Cluster 1



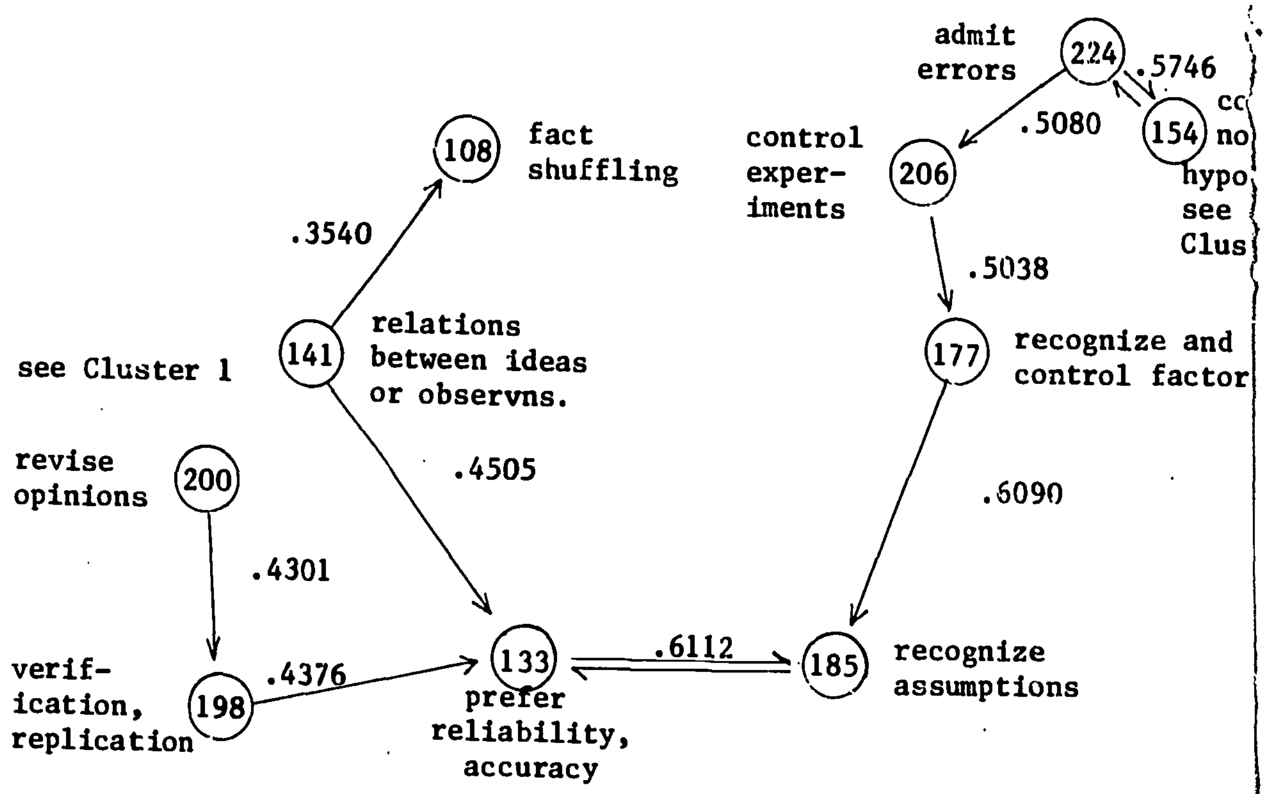
Cluster 1 appears to relate to a commitment of scientists to seek and to report accurately on objective evidence which is the basis for critical analysis of hypotheses and beliefs.

Cluster 2



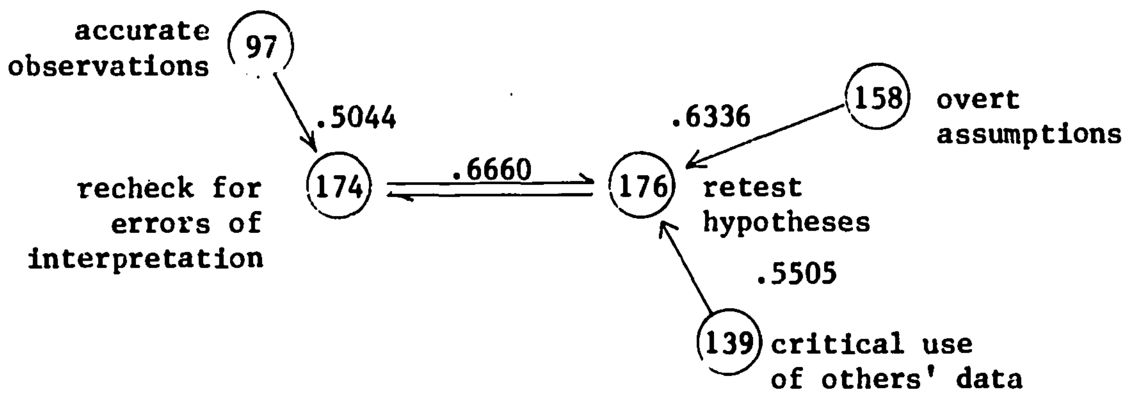
Cluster 2 appears to relate to a preference for initiation of novel and investigative approaches.

Cluster 3



Cluster 3 appears to relate to attitudes of seeking new relationships whilst subjecting them to rigorous experimental control or scrutiny.

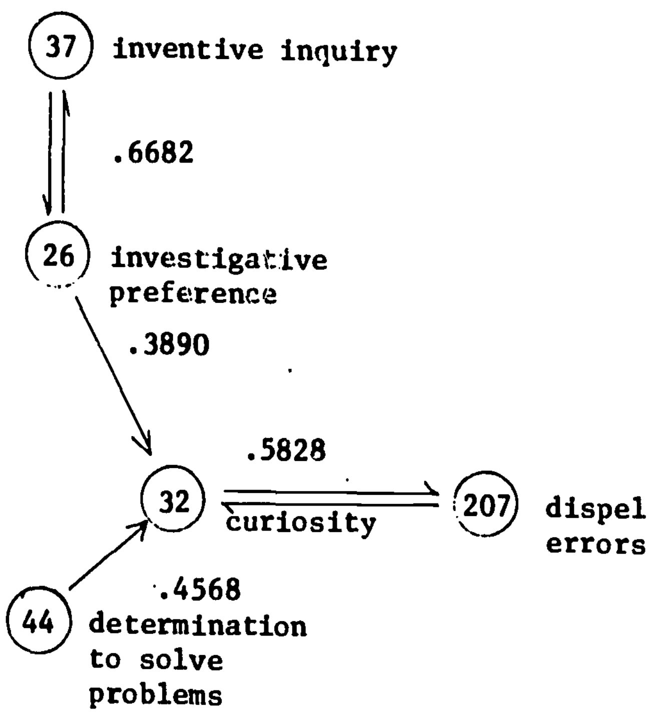
Cluster 4



The attitudes linked in Cluster 4 reflect an emphasis on clarity and accuracy in the use of assumptions, observations and interpretations.

Cluster 5

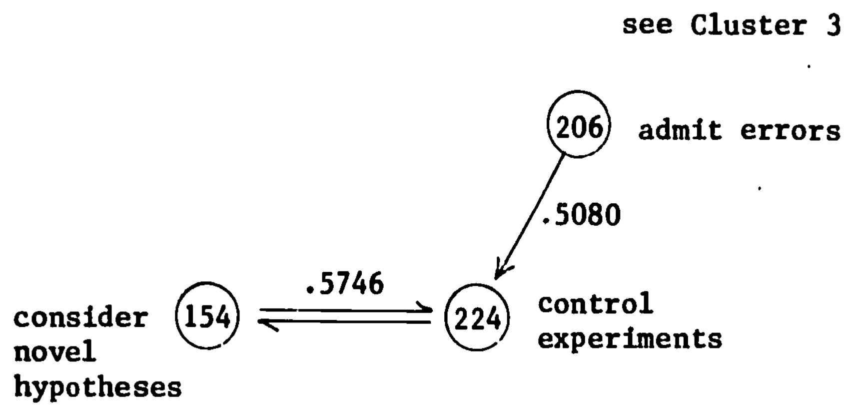
see Cluster 2



The attitudes exhibited in Cluster 5 appear to reflect a thirst for new knowledge and a desire to explore the unknown, as well as a preference for use of inventive techniques for inquiry. However, accuracy and removal of error remain as overriding considerations.

141

Cluster 6



The attitudes portrayed within Cluster 6 suggest that the scientist adopts speculative approaches, but subjects his speculations to controlled experiments.

Discussion and Conclusions

1 Agreement among Australia's most eminent scientists concerning traditionally-accepted statements of scientific attitude was found low.

2 The following scientific attitudes: honest, truthful, sceptical, curious, creative, innovative, inventive;
operating within a known specialist frontier and excluding other considerations, were highly valued.

If we accept this scholarly group as one which can validly identify scientific attitudes, we can extrapolete several implications for curricular objectives:

Implication 1 Students should be encouraged to be critical, and science should not be presented as authoritarian and/or a dogmatic expression of existing beliefs. (cf, Cluster 1).

Implication 2 Open-ended "discovery" approaches should be used in science lessons. (cf, Cluster 2).

Implication 3 Problem-sensing and -solving activities should be used as the bases for laboratory experiences in which control experiments are designed and used. (cf, Cluster 3).

Implication 4 There should be greater opportunity in science curricula for students to identify and express underlying assumptions, to make and record their own observations, and to interpret with clarity and accuracy. (cf, Cluster 4).

Implication 5 There should be opportunity for students to carry out research using "self-designed" techniques and/or apparatus.

142

Implication 6 There should be encouragement of divergent, "non-conforming" ideas, and a tolerance of non-significant results. (cf, Cluster 6).

REFERENCES

ASEP (1970). "Preamble to the Aims of ASEP". *ASEP Newsletter,* No. 2, 4. ASEP, Melbourne.

Bloom, B.S. *et al.* (1955). *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*

Cohen, David (1964). *The Development of an Australian Science Curriculum Model.* Unpublished Ph.D. Thesis, Michigan State University.

Cohen, David and White D.C. (1968). *A New Approach to evaluating attitudes.* Mimeographed, unpublished paper, prepared for the Junior Secondary Science Project (Available from D. Cohen).

Curtis, Francis D. (1932). "A synthesis and evaluation of subject matter topics in general science", in *A Program for Teaching Science,* Thirty-First Yearbook, of the National Society for the Study of Education, Chicago: N.S.S.E.

Edwards, Allen L. (1957). *Techniques of Attitude Scale Construction.* New York: Appleton-Century-Crofts.

Krathwohl, David R. *et al.* (1964). *Taxonomy of Educational Objectives, Handbook 2: Affective Domain.* New York: David McKay.

Noll, Victor H. (1939). *The Teaching of Science in Elementary and Secondary Schools.* New York. Longman's Green.

Rowlands, Evan J. (1971). *Measurement of Scientific Attitudes Held by Victorian Secondary School Children in 1968.* Unpublished M.Ed. Thesis, Monash University.

THE CONTRIBUTORS

EFFIE D. BEST B.Sc. (Hons) (Adel.), Ph.D. (Ohio State), has taught junior secondary school science and senior biology in the UK and in South Australia, and worked for a year as information analyst for the ERIC centre for Science and Mathematics Education in the USA. She is one of the editors of *Biological Science: The Web of Life* (Australian BSCS version) and is presently senior lecturer in biology, Adelaide Teachers College.

DAVID COHEN is Senior Lecturer in Education at Macquarie University. From 1951-62 he taught for the Education Department of Victoria. In 1964, he completed his Ph.D. at Michigan State University in the field of curriculum development. He has lectured in science education at the Technical Teachers College, and was Research Officer for the Education Department of Victoria (1966-68). His main interests are in curriculum development and evaluation, science education and educational media. He is the author of a number number of books and published papers, a member of the Editorial Board of the National Science Curriculum Materials Project, and was editor of the *Australian Science Teachers Journal.*

T.W. FIELD Ph.D.,B.A. (U.N.E.), B.Sc., Dip.Ed. (Syd.), has been a lecturer in Science at Armidale Teachers College since 1963. He recently completed a Ph.D. programme in science education at the University of New England. His major teaching interests are in the physical sciences, curriculum studies and the psychological aspects of science education. The latter area has been the focus of his research activities. He is a member of the Commonwealth Science Facilities Standards Committee and the editor of a current junior science book project.

L.D. MACKAY is Senior Lecturer in Education at Monash University. He has written extensively on the attainment of students following traditional and innovative science curricula. He is currently completing a Ph.D. study at Monash.

G.R. MEYER holds degrees in Arts, Science and Education from the University of Sydney and his Doctorate is from the University of London, Institute of Education. He is a Fellow of the Australian College of Education. For many years Dr Meyer taught Zoology and then Biology at Sydney University and between 1960 and 1966 was the Chief Examiner in Intermediate and Leaving Certificate Biology in New South Wales. Since 1967, he has been the Director of the Centre for Advancement of Teaching at Macquarie University. Recent research interests include the evaluation of various educational media in the teaching of Science and the development of techniques for the formative and summative evaluation of new Science curricula. He is also chairman of the board of consultants of the National Science Curriculum Materials Project (NSCM) which is producing multi media learning materials at matriculation levels for all Australian states.

C.N. POWER, Lecturer in Education at the University of Queensland, has currently completed a Ph.D. study on the effects of communication patterns and feedback in science lessons. Dr Power has served as a science teacher in various state high schools in Queensland, as a Research Officer for the Research and Curriculum Branch, State Education Department, Queensland, and as a Lecturer in Education, University of Queensland. At the University, he has continued his interest in micro-teaching, test construction, implementation of curricula, science education and evaluation, and has developed a modified Bellack Scheme for the observation and analysis of classroom behaviours. He is currently preparing various manuscripts for Australian and overseas books.

GREGOR RAMSEY was Assistant Director (Service) in the Australian Science Education Project. His experience covers curriculum development in chemistry in South Australia and a period as Research Associate at The Ohio State University where he worked with the Educational Resources Information Centre for Science Education. Dr Ramsey is former editor of the South Australian Science Teachers Journal and has been active on curriculum committees and in science teaching activities for many years. He is now Principal of Western Teachers College.

JOHN RENTOUL graduated B.Sc., Dip.Ed. University of Melbourne, 1956. Since then he has held a number of teaching posts, for example, eleven years Senior Science and Mathematics in Victorian State Secondary Schools, on year (1960) Physics and Mathematics at Cranleigh College, Surrey, England, and three years Physics and Mathematics, Knox Grammar School, New South WAles. In 1965, he graduated B.Ed. from University of Melbourne, and at present is completing an M.A. Hons. Thesis in Education at Macquarie University, New South Wales.

R.P. TISHER, Reader in Education at the University of Queensland, was, for some time, a science teacher in secondary schools in New South Wales, and a lecturer in Teachers Colleges. He joined the staff of the University of Queensland in 1963. Dr Tisher's research activities have been associated with science education, the effects of teaching strategies and teacher education. At present he is co-ordinator for the Diploma in Education programme at Queensland and is interested in developing micro-teaching techniques for research projects. During 1969-70, he was a Senior Scholar under the Australian-American Foundation Scheme and a Visiting Professor at the Science Education Centre, University of Iowa.

R.T. WHITE was a teacher in Victorian high schools for ten years, and also taught for a year in New Zealand and for another year at the Melbourne Secondary Teachers College. A year as Senior Master and Acting Headmaster of a high school encouraged him to become Physics Executive Officer of the Victorian Universities and Schools Examinations Board in 1967. After two years in that post he began full-time study in the Faculty of Education, Monash University. He became a member of staff in March 1971. He is now a senior lecturer, specializing in applied statistics and experimental design.

ROSITA YOUNG graduated B.Sc. (N.Z.) from the University of Canterbury with majors in Zoology and Botany. After experience in teaching secondary science in New South Wales, she was appointed to the staff of the School of Education, Macquarie University, where she is currently a Senior Tutor and is completing requirements for the degree of M.Sc. (Hons.).